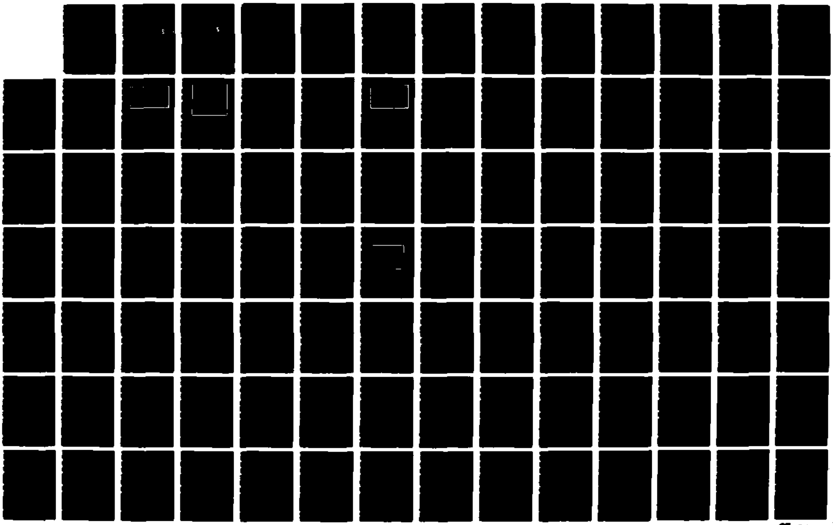
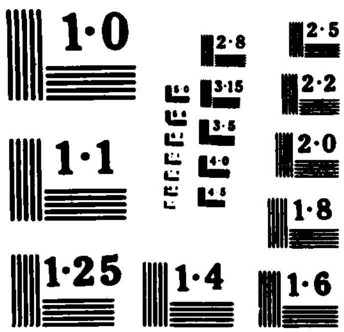
AD-A186 979   ANALYSIS OF THE RELIABILITY OF ROYAL AUSTRALIAN AIR FORCE NON-DESTRUCTIVE (U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF SYST   M CASSIDY   1/1

UNCLASSIFIED   SEP 87 AFIT/GLM/LSMA/87S-11   F/G 14/2   NL



END


1·0
2·8
2·5
3·15
2·2
3·5
1·1
4·0
2·0
4·5
1·8
1·25
1·4
1·6

AD-A186 979

(2)

DTIC FILE COPY

DTIC
ELECTE
DEC 0 3 1987
S D
D

ANALYSIS OF THE RELIABILITY OF
ROYAL AUSTRALIAN AIR FORCE
NON-DESTRUCTIVE INSPECTION

THESIS

Mark Cassidy
Squadron Leader, RAAF

AFIT/GLM/LSMA/87S-11

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

87 11 20 057

(2)

AFIT/GLM/LSMA/87S-11

ANALYSIS OF THE RELIABILITY OF ROYAL AUSTRALIAN AIR FORCE NON-DESTRUCTIVE INSPECTION

THESIS

Mark Cassidy
Squadron Leader, RAAF

AFIT/GLM/LSMA/87S-11

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information is contained therein. Furthermore, the views expressed in the document are those of the author and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.

AFIT/GLM/LSMA/87S-11

ANALYSIS OF THE RELIABILITY OF ROYAL AUSTRALIAN AIR FORCE NON-DESTRUCIVE INSPECTION

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Logistics Management

Mark Cassidy, DipEng(Aero)

Squadron Leader, RAAF

September 1987

Approved for public release; distribution unlimited

Preface

The purpose of this study was to determine a means of establishing the reliability of Royal Australian Air Force (RAAF) Non-Destructive Inspection technicians performing Magnetic Rubber inspections, as well as, to undertake a field survey to derive actual results.

Review of existing literature provided the methodology for analysis and calculation for Non-Destructive Inspection (NDI) reliability. Design and development of an appropriate experimental procedure and NDI Procedure was undertaken and field results obtained. As a result, a 90 percent detection threshold rated at 95 percent confidence for Magnetic Rubber inspection was determined.

In performing the experimentation and writing this thesis I had a great deal of help from others. I am deeply indebted to Mr Derek Olley, Flight Sergeant Kevin Esposito, and other staff of the RAAF Non-Destructive Inspection Standards Laboratory for their knowledge and assistance in developing the NDI Procedure. The participation and constructive criticisms of the RAAF NDI technicians is also gratefully acknowledged. Additionally, I wish to thank my thesis advisor, Captain J. Smith, and thesis reader, Major P. E. Miller, for their continuing patience and assistance in times of need. Finally, I wish to thank my wife Susan who provided much support and motivation to finish.

Mark Cassidy

Table of Contents

List of Figures

List of Tables

AFIT/GLM/LSMA/87S-11

Abstract

The purpose of this research was to establish, via examination of the available literature, an appropriate means of quantifying the reliability of Non-Destructive Inspection (NDI) as practised by the Royal Australian Air Force (RAAF) NDI technicians. Further, actual measurement of this NDI reliability was to be attempted and the correlation, if any, between the NDI technician's reported and measured results and the actual flaw lengths was to be established.

Apart from producing crack size detectibility curves several human factors of the NDI process were to be investigated as part of this research. Influences of personnel variables are considered important. This study was designed to evaluate the effects on NDI reliability on whether or not the technicians, employment has been continuous within the NDI trade; if there is any correlation between experience level and the reliability results obtained; and, the effect of false calls. The effectiveness of reference standards called for by the NDI Procedure was also to be the subject of review.

This study was, unfortunately, constrained by time and lack of resources. Hence, to achieve results the experimental design was modified, with a subsequent effect on the data collection and ability to investigate some of

the research questions.

This study found that the log logistic model was an acceptable Probability of Detection (POD), based on other recent research efforts. However, analysis of reliability results using this model were encouraging, but statistically inconclusive, because of the small sample size available.

Among the recommendations provided are suggestions to improve the experimental procedure, expand the sample size, and continue reliability data collection and analysis to better validate the POD model and answer the research questions made by this study.

ANALYSIS OF THE RELIABILITY OF
ROYAL AUSTRALIAN AIR FORCE
NON-DESTRUCTIVE INSPECTION

I. Introduction

Chapter Overview

This chapter details the general issue, problem statement, research objectives and questions, as well as, the scope, limitations, assumptions, and definitions of this thesis topic on the analysis of the Royal Australian Air Force (RAAF) Non-Destructive Inspection (NDI) reliability.

General Issue

Non-Destructive Inspection (NDI) has, over the past eleven years, evolved into a vital part of the management of aircraft fleet maintenance within the Royal Australian Air Force (RAAF), as well as other aerospace maintenance organizations. Non-Destructive methods of interrogating parts and assemblies for damage caused through normal use, environmental exposure and cyclic fatigue are applied both to detect and assess that damage. Predictions of the extent of damage which can be expected as the result of given aircraft operational roles are now being made on a quantitative basis and NDI is given the responsibility of finding existing flaws. Both major scheduled fleet maintenance at the Depot Level Maintenance (DLM) facilities

and routine maintenance operations at the Operating and Intermediate Levels of Maintenance (OLM and ILM), field levels, employ specific procedures to find flaws before they cause major failures. Additionally, special inspections are performed as a result of unanticipated damage which may appear in-service.

The impact of NDI is felt in operations dealing with both aging aircraft and new weapon systems. The extension of the service life of aging aircraft e.g. the RAAF Mirage, beyond their originally planned use period places the burden of proving structural integrity with NDI. Further, new weapon systems (F111C and F/A-18) have emerged with critical, highly stressed components which are more sensitive to flaws and therefore require periodic NDI to ensure structural integrity throughout their life. In both cases, old and new aircraft, fracture mechanics technology is being used to quantitatively define damage tolerance limits for given flaw sizes and to establish flaw growth rates under specific operating conditions (Stone, 1981:3). This approach to NDI to detect flaws must also be quantified to allow fracture technology to become practical. Consequently, there is a need to evaluate current in-service NDI maintenance capabilities, if possible.

Damage tolerance and structural integrity are treated as the overall guiding concepts in the United States Air Force (USAF) Military Standard on Aircraft Structural Integrity, MIL-STD-1530A. This Standard addresses

requirements such as airframe design, design analysis and development tests, full scale testing, force management, and a force management data package. The possible existence of inherent flaws and the probability of their detection is implicit to all five of the Standard requirements. Design and development functions that interrelate with NDI are treated in the USAF Military Specification on Airplane Damage Tolerance Requirements, MIL-A-83444. This specification asserts that flaws are inherent in any material and designs must account for them. The capabilities of NDI to detect flaws of specific dimensions are assumed possible within the production environment and any design or analysis which assumes inherent flaw sizes smaller than those specified in MIL-A-83444. must stage a demonstration of capabilities to detect those smaller flaws. The American Society for Non Destructive Testing (ASNT) has produced a publication "Recommended Practice for Demonstration of Non Destructive Evaluation (NDE) Reliability on Aircraft Production Parts", which provides guidelines for these demonstration programs. Consequently, an NDI detection capability better than that specified will allow a much greater service life with a corresponding increase in inspection periodicity. Figure 1 illustrates the relationship of fracture mechanics concepts and service life.

In-service flaw detectibility assumptions for airframes have also been set forth in MIL-A-83444.

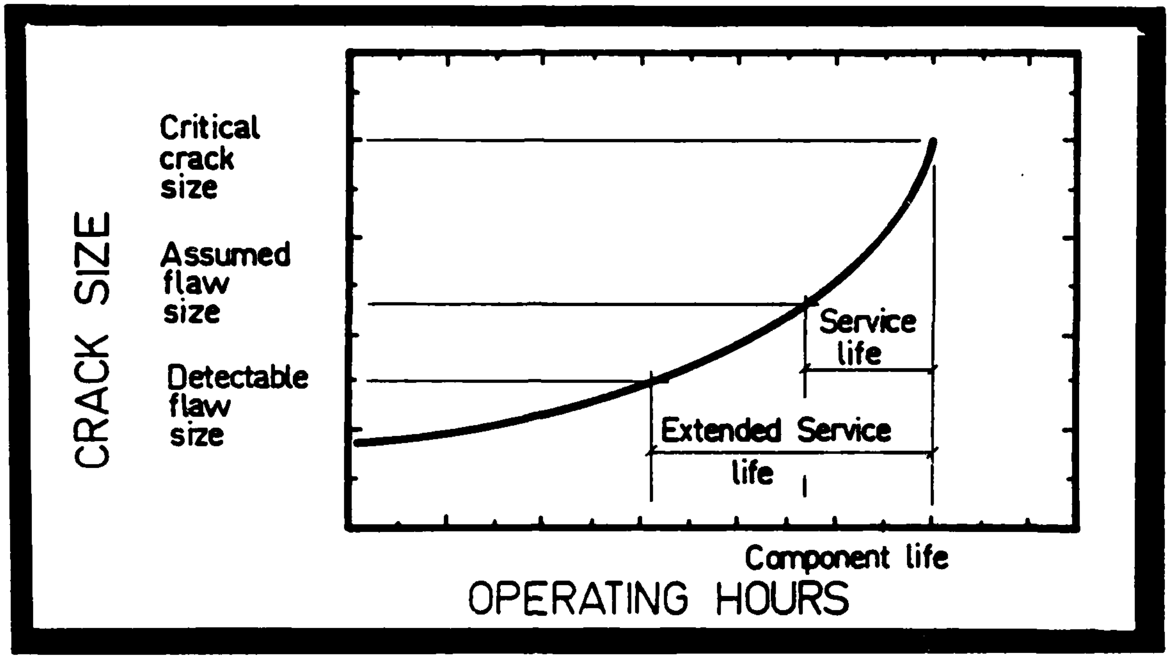


Figure 1. Fracture Mechanics Concept and Service Life Relationship

The NDI Process

Any process of NDI utilizes a system of interacting elements (Figure 2) which comprises the chosen method of NDI, the operator and the specimen (Chin Quan and Scott, 1977:323-354). A further element loosely described as interpretation is often included. During non-destructive inspection, a part is subjected to some type of probing agent, such as radiation, ultrasound, magnetic fields, eddy current or liquid penetrants. A flaw is detected and its size is measured by observing the relative response of the probing agent, which is not always directly related to the

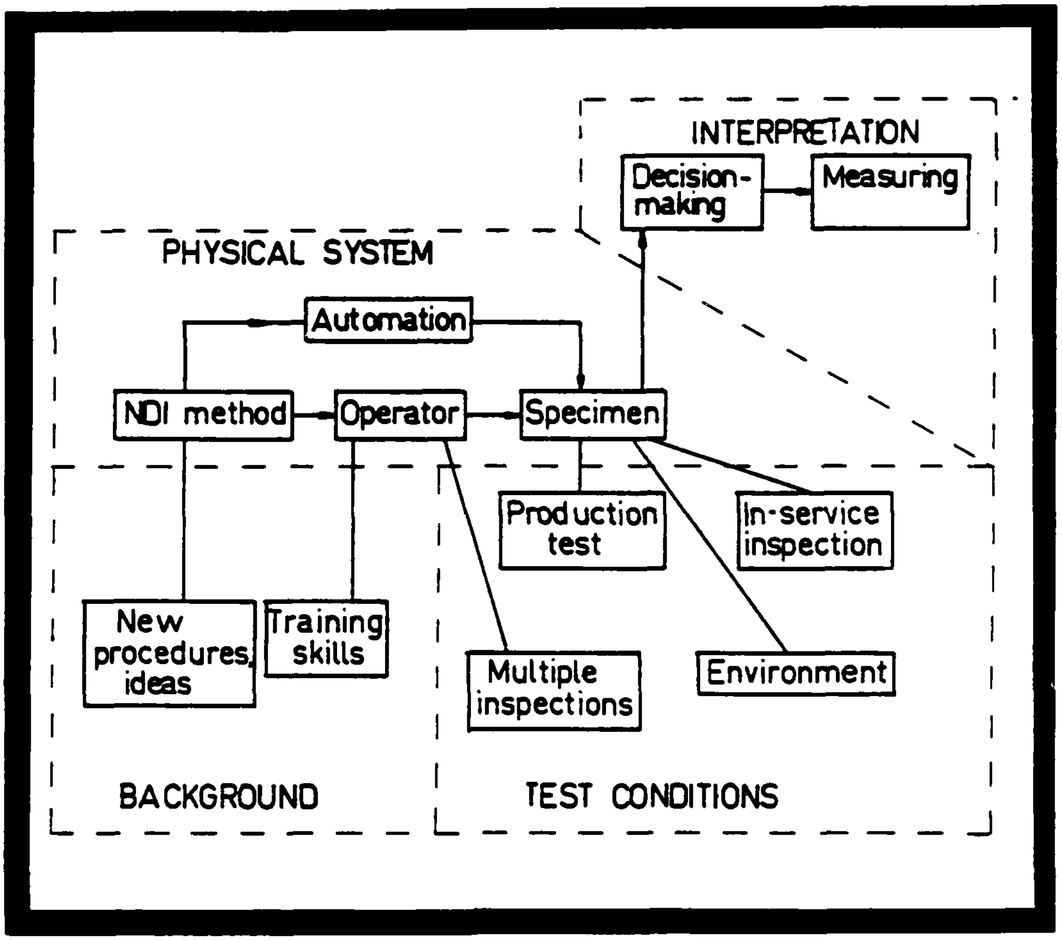


Figure 2. Factors of a NDI System

relative severity of the flaw. Besides the flaw, other characteristics of the part or of the inspection method may affect the response; this introduces uncertainty into an assertion that a flaw has been found or that its actual size has been measured. Additionally, non-destructive inspection is accomplished by human beings. Inherently, no two human beings perform the same repetitive task in an identical manner all the time. This extends to the task of

interpreting inspection results and of making decisions based on these interpretations. Consequently, additional uncertainty is introduced into the inspection process and, in combination with the uncertainty of the probing agent to the flaw, gives rise to the probabilistic nature of the reliability of NDI.

Statement of Problem

Although detection capability probabilities are specified in appropriate Military Specifications (MIL-A-83444, 1974), there is no formal requirement for any demonstration of in-service detection capabilities nor is there any set guidelines for such an in-service demonstration program. Even though, there is no formal requirement, a baseline is still needed to establish experimentally derived values for in-service flaw detection probabilities to ascertain if the RAAF does achieve the assumed MIL-A-83444 goals.

The process of acquiring NDI reliability data is centered around the performance of flaw search tasks on a number of samples with known flaws (Packman, 1976:414). A sufficient number of samples and/or flaw detection attempts are made to establish a statistically adequate volume of data for each selected set of conditions. The demonstration programs are designed to provide evidence that a given flaw size can be detected with a high degree of certainty. A 90% probability of detection at a 95% level of confidence has

been established as the criterion for that high degree of probability (MIL-A-83444, 1976).

Research Objectives

The overriding objective of the research is to determine the existing capability of a selected population of RAAF NDI technicians to detect flaws under field and depot conditions. This will be achieved by establishing flaw detection probabilities for a number of operating and environmental parameters, particularly for the range of aircraft that are meant to comply with the Military Standard and Specification. These detection capabilities will be graphically displayed as detection probabilities relative to flaw size for specific inspection conditions (Figure 3). This information when coupled with appropriate fracture mechanics data such as crack growth rates, allows for quantitative determinations of maintenance inspection intervals (Lund, 1982:2). Additionally, the unique quantified NDI data from the reliability program allows for analyses which point out areas for improvement in operations through the efficient selection of NDI methods, and the optimization of human factors in the management of NDI personnel. Additionally, this program will also allow the RAAF's Non Destructive Inspection Standards Laboratory (NDISL) to better discharge its responsibility to maintain NDI standards by establishing an NDI performance baseline.

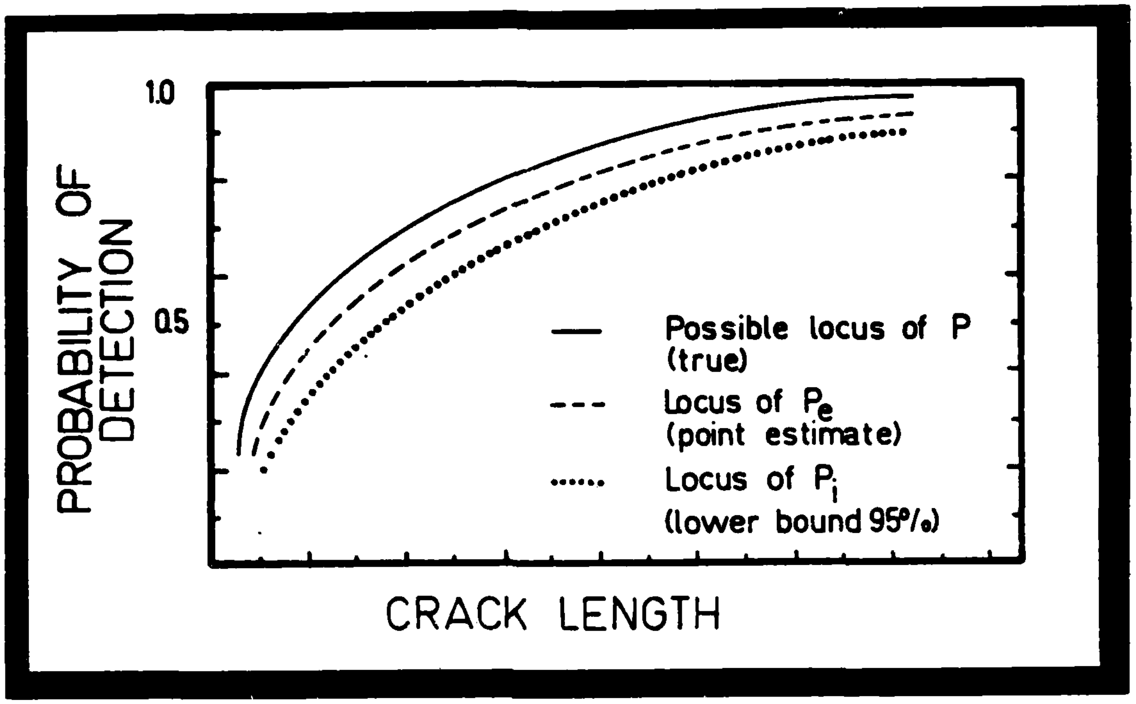


Figure 3. Typical Effect of Crack Length on Probability of NDI Detection

Research Questions

The primary task of the reliability program is to determine results that provide for quantification of flaw detection limits. The following questions will be addressed to answer this research objective:

a. Research Question 1. Determine what the quantifiable flaw detection limits are for the RAAF NDI population at RAAF Base Amberley.

b. Research Question 2. Establish what correlation there is, if any, between the NDI technician's

reported and measured results and the actual flaw lengths.

Although crack size/detectibility curves form the basis for the measure of the RAAF NDI reliability, they are still only a preliminary measure, as the NDI Technicians' inferred performance needs to be qualified by other factors. For the NDI system described in Figure 2, there are numerous variables. Factors or variables that can influence the NDI reliability results are to be considered and measured as part of the reliability program. The effect and change of some of these variables have significant impact on the reliability program results. Unlike the USAF, who have a dedicated trade group for NDI personnel, the RAAF trains its basic aircraft trade musterings (ie Airframe and Engine Fitters) to perform NDI functions. Hence, influences on the personnel variables are considered very important in the research study to optimize human factors important to successful NDI.

a. Research Question 3. The effect of whether or not the reliability results are dependent on whether the technician's employment has been continuous within the NDI trade is to be determined.

b. Research Question 4. Determine if there is any correlation between the experience level (as measured by the NDI technician's years of

employment in that field) and the results obtained.

c. Research Question 5. Establish the effectiveness of reference standards, which are employed to standardize and calibrate NDI Procedures.

d. Research Question 6. Determine if the frequency of false-alarm calls has any significant effect on the reliability.

Scope of Study

Because of time constraints, this thesis research cannot evaluate the reliability of all the NDI methods, nor can an evaluation be performed on all RAAF NDI technicians. Hence, only the most widely used method, Magnetic Rubber, will be the subject of review. Additionally, only the NDI technicians that regularly use this method will be the subject of investigation. Magnetic Rubber is a combination of a room temperature curing silicon rubber solution which contains very fine ferro magnetic particles. The use of a catalyst and cure stabilizer ensures an even solidification process allowing controlled migration of the ferro particles under the influence of a magnetic field. The products used in this study will be Magnetic Rubber (Part Number MR502) manufactured by Dynamold Incorporated of Texas, USA. This product was specifically developed for use on the F111

aircraft. Nevertheless, the reliability program approach is just as applicable to all the other NDI methods and will form the basis for an on-going reliability program for evaluation of all commonly employed NDI methods.

Limitations

Although the RAAF have extended the use of Magnetic Rubber to aircraft types other than the F111C, this reliability program will be restricted only to RAAF Base Amberley NDI personnel participation. RAAF Base Amberley is the home for the Australian F111C Squadrons; and, as well, is the location of the Non-Destructive Inspection Standards Laboratory (NDISL). This approach is reasonable since the majority of NDI technicians are employed within RAAF Base Amberley, and the requirements for establishing detection probabilities are essentially F111C related, at this point in time. Further, only personnel qualified and certified, in accordance with MIL-STD-410D as Level 1 or 2 NDI technicians, will be tested. This has the effect of eliminating non practicing NDI technicians, and allows trainee NDI technicians currently undertaking on-the-job training, before being examined and certified as proficient in that method, to be included in this study.

Assumptions

Previously, it was stated that the size of a flaw is measured by observing the relative response of the NDI

probing agent, which is not necessarily directly related to the severity or expected flaw characteristic (eg flaw length). However, whenever any flaws are measured as part of this reliability program, using the indications left on the Magnetic Rubber replica, the measurement taken from the Rubber is to be assumed to be the same length as that of the actual flaw detected. Any inconsistencies occurring as a result of this assumption will be minor, as the primary aim of this study is the detection of flaws. Measurement of flaw characteristics is of secondary importance to the thesis objectives.

Due to the unavailability and inappropriateness of suitable reliability test pieces, such as actual aircraft components, this reliability study will use test coupons. But, the relationship of the results from test coupons and actual aircraft airframe structures is unclear (Lewis, Dodd and other, 1978:6-5). Accordingly, this research will assume, for the time being, that the results obtained using test coupons are directly related to the detection limits for actual aircraft inspections.

Another assumption made is that the NDI Procedure, which details the inspection process to be used, allows several test coupons to be tested at the same time. While this is the practice for real life field inspections, from a statistical point of view, the results obtained are not from completely random independent samples. The main concern is that a confounding variable, such as Magnetic Rubber batch

sensitivity, may impact the results. However, a reference standard to standardize the inspection process, is to be incorporated into the study to negate any potential uncontrolled influences.

Definitions

Throughout this thesis various technical terms relating to NDI are used that need to be defined to ensure a consistent approach and universal understanding. The following definitions are supplied as a result.

Non-destructive Inspection (NDI). NDI is a standard maintenance procedure used in the RAAF. NDI refers to the various means which can be employed to examine components without damaging them in any manner (Engineering Design Handbook, 1976:8-1). Methods of NDI in current use in the RAAF or readily available to technical staff are detailed in DI(AF)AAP7002.008-1. The terms evaluation and testing are sometimes used in lieu of inspection, because of philosophical differences in application of the NDI process.

Level 2 NDI Technician. There are three levels of technician qualification. The Level 2 person is able to direct and carry out inspections and certify results and to conduct and grade examinations, prepared by a higher qualified NDI personnel, for qualifying subordinate NDI technicians (MIL-STD-410D, 1978).

Level 1 NDI Technician. The Level 1 person is a trainee on a formal NDI course, or a person who has

completed the formal NDI training course but is still undergoing on-the-job training before qualification and certification as a Level 2 NDI technician. A Level 1 technician can perform all or part of NDI tasks while under the supervision of a higher certified NDI technician (MIL-STD-410D, 1978).

<u>Probability of Detection (POD).</u> The probability of flaw detection is the probability that, using a given inspection procedure, a trained inspector will detect a flaw if it exists (Packman, 1976:415). A probability of detection of 90% implies that for every 100 flawed parts that are inspected at least 90 of them would be identified as containing flaws, and that no more than 10 parts containing flaws would be identified as being free of flaws.

<u>Degree of Confidence.</u> The degree of confidence in the probability of detection refers to the ability to estimate from a limited sample the probability of detection representative of large-scale inspection (Packman, 1976:415). For example, a probability of detection of 90% implies that at least 90 out of 100 flawed parts would be identified each time a sample from a given population is inspected. The level of confidence refers to the probability that this implication is valid. Thus, 90% detection probability with 95% confidence means that there is a 5% probability that 90% is an overestimation of the true (unknown) detection probability.

Chapter Summary

This chapter has defined the problem to be explored as this thesis research, as well as, setting the constraints for this study. Assumptions and definitions are provided to better establish the problem bounds. The following chapter will detail the literature review undertaken to determine an appropriate research methodology.

II. Literature Review

Chapter Overview

This chapter details the general theory, specific methods available, prior reliability efforts, and evaluation methods considerations which impact this thesis topic on the analysis of the Royal Australian Air Force (RAAF) Non-Destructive Inspection (NDI) reliability.

General Theory

According to Packman in his article in the ASM Metal Handbook, the probability of flaw detection is defined as the probability that, using a given inspection procedure, a trained inspector will detect a flaw if it exists (Packman, 1976:415). This probability of detection can be determined by experimentally observing the number of times a non-destructive inspection procedure can reveal flaws in a sample of parts known to contain flaws.

The true probability of detection, P, is an unknown quantity that cannot be determined without making an infinite number of inspections. The main aim of statistical evaluation of the reliability of non-destructive inspection is to estimate P with as few trial inspections as possible. This is done by defining a lower limit (lower bound) for the range of values that is expected to contain the true probability of detection. Figure 3 graphically represents this situation, without entering into mathematical aspects.

The top curve is a possible locus for values of P, the true probability. The middle curve represents the locus for the values of P_e, the point estimate of the true probability. The point estimate is independent of sample size and is simply the experimentally determined ratio of the number of flaws detected to the number of existing flaws. The bottom curve is the locus of P_l, the values known to a specified lower-bound (one sided) confidence limit. These are calculated values whose position relative to P_e depends on sample size. Each value of P_l is intended to be a conservative estimate of the true probability, in that it is expected to be on a curve that falls on or below the curve for P (Packman, 1976:416).

Initial Reliability Efforts

Historically, NDI reliability work in the aerospace industry began with the recognized need for such data to interplay with the evolving fracture mechanics analyses in the late 1960s. The first detailed investigation was conducted by Packman for the USAF Materials Laboratory in 1969 (Packman, 1976:415). The objective was to measure flaw detectibility for aircraft production parts. A number of programs have been sponsored by both the USAF (Lewis, Dodd and others, 1978) and NASA (Rummel, 1974) since that time, with a diversity of specimen configurations and flaw types used in flaw detection tasks. Chin Quan and Scott of Aeronautical Research Laboratories (ARL), Melbourne,

Australia, in their article on Operator Performance and Reliability in NDI give an extensive history of the serious attempts, including their own, to compare the reliability of NDI (Chin Quan and Scott, 1977:323-354).

Two reliability programs are of special interest. The first NDI reliability assessment on a built up structure, as contrasted to production part configuration, was conducted as internal research at the Lockheed - Georgia Company beginning in 1971 (Lewis, Dodd and others, 1978:1-3). The results of this study by Sproat and Dodd showed that fatigue crack detection probabilities in assembled structures would be generally lower than for comparable flaws in parts or specimens. However, this study did not elaborate as to any potential relationship between the two detection probabilities i.e. assembled structures versus specimens. Two conclusions at the time were:

1. Fatigue crack detection probability on structure, by single application of ultrasound and eddy current NDI procedures, is significantly lower than that normally assumed for most fail-safe and slow crack growth airframe designs.
2. Redundant inspections using multiple applications of one procedure or a mixture of procedures can be used to yield the detection levels required for slow crack growth and fail-safe structures (Lewis, Dodd and others, 1978:1-3).

However, of more fundamental importance was the results

of a four year USAF Logistic Command program (Lewis, Dodd and others, 1978) to determine the reliability of USAF non-destructive inspection capability. This report was completed in 1978 and is usually known as 'Have Cracks Will Travel' program. Actual aircraft structural samples containing fatigue damage were transported to 21 different Air Force bases and depots where about 300 technicians used a variety of NDI methods on the samples.

The measured probabilities of detecting fatigue cracks in built up structures from the "Have Crack" program were not as high as required or expected. The primary source of variance between the individual technicians was in the human factors area. The specific variables, under evaluation, of formal education, age, classification skill level, NDI experience, and NDI training were each analyzed and proved to have only minimal influence on the resulting NDI performance (Lewis, Dodd and others, 1978:12-3).

Apart from an engine component NDI study (Rummel, 1984), there does not appear to be any further detailed and comprehensive studies conducted since then which have published their results, although there has been a number of evaluation trials by the USAF for new products or procedures. While previous studies had cited human factors variation as the primary factor in NDI performance, this aircraft engine NDI study concluded that the overall NDI performance level was dependent on the adequacy of the NDI engineering aand acceptance criteria definition; the NDI

materials; equipment; processes (methodologies); and human skills applied to the task (Rummel, 1984:213). Basically, if the NDI engineering, materials, equipment and processes are not under control, then the influence of the human element is not important.

Research, since then has concentrated on statistical analysis of reliability data. Emphasis has been on validating the various NDI reliability experiments available. There are three categories of experiments which have been used to evaluate NDI systems reliability; namely, 1) demonstration of a capability at one crack length, 2) estimation of the probability of detection (POD) function and associated confidence bounds through single inspections of cracks covering a range of lengths, and 3) estimation of the POD function and confidence bounds through multiple inspections of cracks covering a range of lengths (Berens and Hovey, 1981:5). Analysis of data from the first two categories have generally been based only on binomial distribution theory. While the last experiment category data have been analyzed by regression analyses. This category of experiment resulted from the 'Have Cracks Will Travel' program. But, no rationale for the functional model used, other than goodness of fit, was presented for this approach. However Berens and Hovey recently used data from this "Have Cracks" study to develop a functional model and validate this approach to characterize the probability of detection function using regression analysis (Berens and

Hovey, 1984:139).

Details of the three types of experiments and the analysis methods used are presented in the Packman and Berens References.

Estimation of the POD Function with Multiple Observations per Crack - "Have Cracks Will Travel"

Results from the "Have Cracks Will Travel" program clearly illustrate that not all cracks of the same length have the same detection probability when subjected to independent inspections by different inspectors (Berens, 1981:16) This method of collecting data using multiple observations per crack yields an estimate of a detection probability for each individual crack.

To analyze the data collected from this category of NDI experiment, a regression analysis is performed in which a model curve is fit to data points and a lower confidence limit is placed on the regression equation. In the "Have Cracks Will Travel" program, the model selected as providing the best fit is given by

$$POD(a) = \exp\left[-\alpha . a^{(1-\beta)}\right] \qquad (1)$$

where the parameters are estimated by a linear regression on transformations of the crack lengths and observed probabilities of detection. This functional model was subsequently called the Lockheed model (Berens, 1981:16).

This "Have Cracks" study took the viewpoint that the POD at a particular crack length was a low percentile of the

distribution of detection probabilities at the crack length. To calculate a lower confidence limit on the POD, a confidence bound on the population of detection probabilities was used (Berens, 1981:16). This approach is different from that of traditional use for lower bound limits. It was established later by Berens that the POD is actually the mean of the detection probability distribution. Hence, the POD confidence limits should be placed on the mean regression line and not on the total population of detection probabilities (Berens, 1983:24-26).

<u>Evaluation of "Have Cracks Will Travel" Methodology</u>

Berens, in his study, investigated various functional models for the POD curve. The "Have Cracks" data being representative of field inspection capabilities was selected for the study to determine an acceptable model for the POD(a) function. Three criteria were established for the definition of "acceptable"; namely, (1) goodness of fit, (2) normality of deviations from fit, and (3) equality of variance of deviations from fit for all crack lengths (Berens, 1981:21). These are standard statistical measures for regression analysis modeling. The latter two criteria are necessary statistical assumptions for the validity of confidence limits derived from regression analyses.

Seven functional forms were investigated in the selection of an acceptable model (Berens, 1981:22). Potential POD functions are listed at Table I. The Lockheed

model (equation 1) was derived during the original analysis of the "Have Cracks" data. The Weibull model was selected since it is a generally accepted model and is a variation of the Lockheed model. The other five models were selected because they have been found to be useful in analogous problems in the field of bioassay.

Regression analyses were used to fit all seven models to the "Have Cracks" data. The detection probabilities, p_i, and the crack lengths, a_i, for each crack were transformed to y_i and x_i in accordance with the transformations of Table I. The transformed x and y variables were then used in a linear regression analysis of the form

$$y_i = A + Bx_i + e_i \qquad (2)$$

(Berens, 1981:23)

For all seven model, B is the estimate of b and, depending on the model, either A or exp(A) is the estimate of a. The deviations of the transformed observations from the regression equation, i.e., were analyzed to test the applicability of each model with respect to the three acceptability criteria.

The probit, log odds-linear and arcsine models were based on models that did not transform the crack length scale. None of these models were found to provide adequate goodness of fit in the sense that their patterns of deviations were

TABLE I

Potential POD Functions

Name	Functional Form	Transformation
Lockheed	$P(a)=e^{-\alpha(a)^{1-\beta}}$	$y=\ln(-\ln(p)/a)$, $x=-\ln(a)$
Weibull	$P(a)=1-e^{-\alpha(a)^{\beta}}$	$y=\ln(-\ln(1-p))$, $x=\ln(a)$
Probit	$P(a)=\phi(\alpha + \beta.a)^{*}$	$y=\mathrm{PROBIT}(p)$, $x=a$
Log Probit	$P(a)=\phi(\alpha + \beta.\ln(a))^{*}$	$y=\mathrm{PROBIT}(p)$, $x=\ln(a)$
Log Odds -linear scale	$P(a)=\dfrac{e^{\alpha+\beta a}}{1+e^{\alpha+\beta a}}$	$y=\ln(p/(1-p))$, $x=a$
Log Odds -log scale	$P(a)=\dfrac{\alpha a^{\beta}}{1+\alpha.a^{\beta}}$	$y=\ln(p/(1-p))$, $x=\ln(a)$
Arcsine	$P(a)=\sin^{2}(\alpha + \beta.a)$ $0\le a\le(\pi-2\alpha)/\beta$	$y=\mathrm{arcsine}(\sqrt{p})$, $x=a$

*

$$\phi(x)=\int_{\infty}^{x}\frac{1}{\sqrt{2\pi}}\,e^{-1/2y^{2}}\,dy$$

(Berens,1981:22)

not randomly distributed about the model over the entire crack length (i.e. they were inconsistent with the linear model of the equation (2)). These models were rejected on this basis. The other models generally provided an adequate

fit to the observed data (Berens, 1981:23).

Berens then used the Bartlett's test and Shapiro-Wilks W test to evaluate the equality of variance and normality, respectively, of the deviations from the regression equations. The variance of the deviation from the log odds-log scale and log normal models was constant; however equality of variance was rejected for the Weibull and Lockheed models (Berens, 1981:24).

The log odds-log scale (or log logistic) model was consistent with the assumptions of normality of deviations. None of the other models performed as consistently. Therefore, Berens concluded this model provided an adequate fit for the "Have Cracks" data (Berens, 1981:24). Further, the log logistic distribution always provided the lowest estimate of POD in the tails of the function. This indicates that the log logistic function represents a conservative choice for the POD model which is a valid feature when dealing with structural safety. Additional Berens research concluded that, given an acceptable model for the regression function, the regression estimates of NDI capability are superior to those derived using binomial distribution theory (Berens, 1981:74). The regression estimates are closer to the true POD, exhibit less scatter in the distribution of the estimates, and, contrary to the binomial methods, always provide an estimate of the desired confidence limit.

Influence of Experiment Design

Berens also conducted research on the effect of the design of the experiment, using simulations based on POD and flaw characteristics from the "Have Cracks" data. In this instance the design of the experiment referred to the number and distribution of the flaw sizes used in the reliability demonstration program.

The basic effect of sample size on the POD parameter estimates was evaluated via simulation. The range of parameter estimates and therefore the scatter, decreases with sample size. The standard deviations of the POD estimates are theoretically proportional to $1/\sqrt{n}$ (where n is the sample size) and the Berens simulation results generally agreed with this reduction. Additionally, the influence on the POD function estimates as a function of the standard deviation of the flaw sizes used in the simulation.

As a result of the Berens research the following conclusions were drawn

> The foremost consideration in designing an NDI reliability demonstration should include flaw sizes that span the full range of POD values from 0 to 1. Fairly stable distributions..were obtained when flaw sizes spanned the POD function and the sample size was 30 or larger [Berens, 1981:80].

Treatment of Special POD Cases

A problem in the use of regression analysis arises when the observed proportion of detected cracks at a crack length is zero or one. In either of these cases, the most useful

transformations can be undefined (Berens, 1981:29) To overcome this problem several solutions are available. The "Have Cracks" program substituted the values of 0.01 and 0.999 for 0 and 1, respectively (Berens, 1981:29).

Berens proposed a more acceptable alternative, which was to use a different estimator for the detection probability. The usual (maximum likelihood) estimator for the detection probability taken as

$$\hat{p} = i/n \tag{3}$$

where i is the number of detections and n is the number of specimens with the crack of the fixed length. Another estimate of the proportion with acceptable statistical properties is the mean estimate

$$\bar{p} = i/(n+1) \tag{4}$$

This method depends on n being large i.e. 30 or greater.

False Call Treatment

The inspection process constitutes an exercise in conditional probability as opposed to joint probability due to the interdependence of inspection stimuli and inspection responses (Rummel, 1984:8). The following schematic presentation shows such interdependence:

	STIMULI POSITIVE a	STIMULI NEGATIVE n
RESPONSE POSITIVE A	TP NO ERROR	FP TYPE II ERROR
RESPONSE NEGATIVE N	FN TYPE I ERROR	TN NO ERROR

(Rummel, 1984:8)

The outcome of the inspection may be:

1. True Positive (TP) call - no error condition,
2. False Positive (FP) call - Type II error condition (finding a flaw when none exists),
3. False Negative (FN) call - Type I error condition (failure to find a flaw when one is present), and
4. True Negative (TN) call - no error condition.

The Probability of False Alarms (POFA) can be expressed as:

$$POFA = \frac{FP}{TN + FP} \quad \text{or} \quad \frac{\text{total false alarms}}{\text{opportunities for false alarms}} \tag{5}$$

(Rummel, 1984:8)

Rummel classified the errors in performance by skilled operators as:

1. Systematic Error, which are consistent offsets from the ideal performance,
2. Errors in Precision, which are consistent but random variations in performance, and
3. Sporadic Errors, which are occasional occurrences varying significantly from the predictable performance (Rummel, 1984:223).

During the Aircraft Engine Reliability study (Rummel, 1984), sporadic errors in detection at large flaw sizes resulted in data scatter at the large flaw sizes, and was usually due to sporadic human error. Such errors were attributed to drowsiness, lack of interest, lack of motivation, fatigue, boredom, monotony, etc (Rummel, 1984:224). These error could be minimized by attention to the factors responsible for the problem, and by redundant

inspections.

Errors in precision were indicated by data scatter at the transition region of the POD curve (Rummel, 1984:224). Errors in precision can be caused by slight variations in processing, by inexperience of the NDI technician or a shift in decision criteria (usually caused by a lack of confidence). Experience, expert skill development and well defined and recognized acceptance/rejection criteria will minimize this error mode (Rummel, 1984:224).

Systematic errors are indicated by a shift in the threshold point on the POD curve when inspection is performed on identical components by two different operators. Differences in performance may be due to differences in skill and/or decision criteria by the NDI technicians. Proper training and direction regarding decision criteria can reduce systematic errors between inspections and between operators (Rummel, 1984:226).

Chapter Summary

This chapter presented a synopsis of general reliability theory, and a history of the initial research efforts in this field. Greater detail was given on the "Have Cracks" reliability program and the subsequent

statistical analysis by Berens of the methodology employed in that program. This chapter provides the framework from which the next chapter on methodology is based.

III. Methodology

Chapter Overview

This chapter details the population, sample size, experimental procedure, field study inspection instrument, as well as, the data processing and analysis, and statistical analysis requirements for evaluating the RAAF NDI reliability.

Population

The specialist nature of NDI duties and the sporadic workload makes pooling of NDI resources and facilities at particular locations essential. Where a significant workload exists at a base, an NDI Section is formed to provide services for all units on that base. Within the RAAF, virtually all bases have centralized NDI Sections. The total number of RAAF Level 1 and 2 NDI technicians is 54. They are employed at eight bases; but, more than half (33) are located at RAAF Base Amberley in Queensland. This research is being restricted to RAAF trained, qualified, and certified Level 1 and 2 NDI technicians at RAAF Base Amberley, presently performing NDI tasks. These technicians have all been selected from the Airframe and Engine trades, and have completed an NDI Technician Course before beginning NDI duties. Experience levels range from trainees and recently certified personnel to members with 11 years

practical application of NDI. Rank levels vary from Leading Aircraftman to Flight Sergeant.

Sample Size

As discussed in Chapter II, consistent results are obtained when a relatively large sample size is used. A sample size of 30 or larger was recommended. Sample sizes as low as five to 10 were still producing results; but, at least 10 inspections should be performed (Berens, 1984:139). Due to the geographic distances and the constraints of time, a pre-survey to establish the population statistics for a number of research question parameters (e.g. the number of members not continuously employed within the NDI field) was not possible. Hence, because of the small population for this experiment, the sample sizes obtained are likely to fall below the recommended limits. Ranges for the sample size are expected within five to 10. As a consequence, a number of Chapter I research questions may not be able to be conclusively evaluated.

Experimental Procedure

The design of the demonstration program is extremely important in ensuring validity of results. Also demonstration programs can be rather expensive in terms of resources. In general, the sequence of steps in a reliability demonstration program are:

1. Design a specimen having the size, shape and

surface finish that approximately represents the actual component to be inspected.

2. Decide on the type of flaw for which the inspection procedure is to be evaluated (eg fatigue cracks, inclusions). Flaw location and orientation should be specified. Introduce these flaws into the test specimen.
3. Select the flaw characteristic (such as length, depth, or area) for which the inspection program is to be evaluated.
4. Decide on the flaw-size range to be investigated. On the basis of statistical requirements determine the minimum number of observations needed for qualification of the inspection procedure.
5. Prepare and identify the required number of flawed specimens and at least an equal number of flaw free specimens as control.
6. Develop a complete, detailed, clear and unambiguous inspection procedure.
7. Randomly mix specimens containing flaws of different sizes with flaw-free control specimens and inspect in strict accordance with the written procedure.
8. Upon completion of the inspection program, reconfirm the flaw characteristic.
9. Collate the data and graph the results.
10. Decide on the certification for the type of flaw

used in the demonstration program.

11. Repeat, if necessary (Packman, 1976:416).

For the NDI described previously, there are numerous variables. For the development of a realistic reliability program, a number of factors will be classified as controlled and uncontrolled variables. The controlled variables need to be specified (Lewis, Dodd and others, 1978:2-1 to 2-4) and include the NDI method (eg Magnetic Rubber), the use of a detailed inspection procedure and calibration standards, specimen configuration (ie position, flaw size population, specimen complexity and randomness, flaw characteristics). The effect and change in these controlled variables are the principal subjects for evaluation. The uncontrolled variables including environment, human physiological response, attitude (psychological), personnel and disruptive factors, are the factors which the researcher needs to be wary of as the reliability program is unable to quantify their influence on the final result.

In this particular reliability study nine D6AC Ultra High Strength steel test coupons have been developed, and flaws, fatigue cracks, induced (Figure 4). They were flawed under the influence of high cycle low stress within a corrosive environment i.e. corrosion induced fatigue cracks. This steel has been heat treated to 260 - 280 ksi. The number of flaws in these test pieces total 236, ranging in size from 0.040mm (less than two thousandths) to 5.7mm (225

thousandths of an inch). A further seven test coupons were used as placeboes (i.e. not flawed). The flaw characteristic, which will be the subject of inspection and measurement, is the crack surface length. Certification of the results will be based on the 95% confidence for 90% detection probability.

Table II shows details on the flawed D6AC test coupons used and some of the flaw characteristics.

TABLE II

D6AC Test Coupon Particulars

Coupon ID	Hole #	# of flaws	Flaw Range (mm) Minimum	Flaw Range (mm) Maximum
4	4	8	0.1	5.7
53	1	4	0.538	0.968
	3	11	0.313	1.677
	4	22	0.047	3.594
59	1	8	0.381	1.815
	3	6	0.133	1.600
	4	10	0.388	1.647
82	1	2	0.410	0.910
	3	4	0.100	0.870
	4	11	0.040	0.430
145	3	10	0.163	2.237
	4	7	0.893	1.964
160	3	12	0.100	1.240
	4	11	0.230	1.740
175	1	5	0.087	0.540
	3	4	0.270	1.723
	4	8	0.135	3.041
272	1	25	0.080	1.842
	2	9	0.087	0.301
	3	7	0.180	0.703
	4	7	0.160	0.522
N	2	9	0.070	0.710
	3	19	0.070	4.170
	4	17	0.130	0.880

Apart from the D6AC test coupons listed in Table II, the following unflawed test pieces, or placeboes, were part of the experimental controls: 43,64,71,91,120,131, and 136.

Appendix A details the experimental procedure used for this field study.

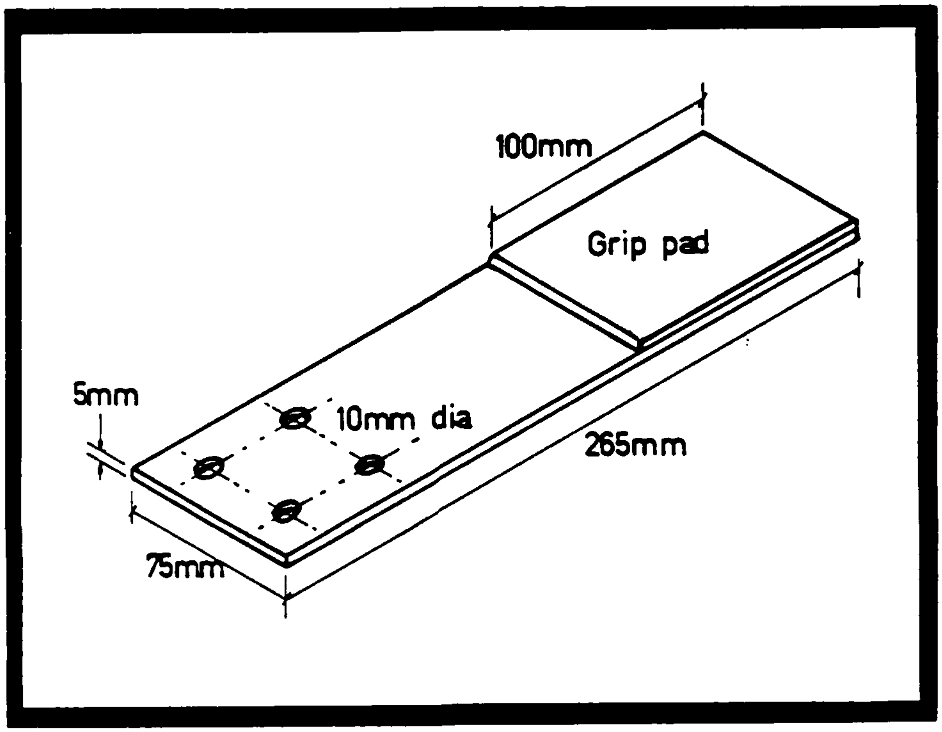


Figure 4. Test Coupon - D6AC Steel

Field Study Inspection Instrument

Appendix B details the NDI Procedure to be followed by the RAAF NDI technicians undertaking this field study inspection. Although the NDI Procedure does not allow any flexibility or interpretation in the inspection process, it does provide for a number of Magnetic Rubber types to be tested, and subsequently evaluated. This Procedure will be

used, with and without the nominated reference standard, to determine the effect of reference standards.

Validation of NDI Procedure

The NDI Procedure was written in accordance with RAAF Specification (Engineering) P30 which details the -36 topic manual format for RAAF NDI Procedures. The Procedure was validated by NDISL staff before assignment to technicians, under test conditions. The Procedure produced significant indications on all cataloged flaws. Some other indications, however, could be classified as ambiguous due to the surface condition of the hole walls e.g. mechanical scoring.

Data Processing

Storage, analysis and graphing of all statistical data is to be done on a micro based computer system for ease of use within the RAAF NDI operating environment. The computer to be used will be an IBM XT clone, a Zenith Data Systems model 158, and the software will be commercially available Data Base Management System (DBMS) and statistical analysis packages. The Ashton-Tate Corporation product, DBASE III Plus, and the Lotus Development Corporation software, Lotus 1-2-3, are employed to provide the appropriate data storage, analysis and report generation. Lincoln Systems Corporation's Interactive Statistical Programs are used to determine statistical significance of relationships. The minimum hardware and software configuration required is an

IBM PC, XT, or AT, or other 100 percent IBM compatible computer with a monochrome or colour monitor. The system should have at least 256K RAM of memory (512K RAM or more is suggested), an operating system of DOS 2.00 or higher, two 360K diskette drives or one 360K drive and a hard disk, and a printer with at least 80 column capability.

Statistical Analysis

The analysis technique for estimating the probability of detection (POD) function, for NDI results recorded in a pass/fail form, depends on the data type. For this study, since there are multiple inspections of each flaw, and a large number of flaws, a regression analysis can be used to estimate the parameters of the probability of detection model (Berens and Hovey, 1984).

The analyses are based on a log logistic function (Berens and Hovey, 1981:139) which is made to fit to data points. The model selected as providing the best fit is given by the functional form shown at Table I. However, a direct analysis of the model when expressed in the form given in this table is very complicated. The analysis can be simplified by using the re-parameterized model

$$\text{POD (a)} = \frac{\exp(\alpha + \beta.\ln(a))}{1 + \exp(\alpha + \beta.\ln(a))} \qquad (6)$$

(Berens, 1984:145)

where the parameters are estimated by a linear regression on transformations of the crack lengths and observed

probabilities of detection. Confidence limits are then calculated on the mean regression line.

This log logistic (or log odds) model comes from the logarithm of the odds $(p/(1-p))$ (log odds) transformation. The log odds transformation converts equation (6) to

$$\ln\left(\frac{POD(a)}{1-POD(a)}\right) = \alpha + \beta . \ln(a) \tag{7}$$

which is linear in the transformed variables

$$Y(a) = \ln\left(\frac{POD(a)}{1-POD(a)}\right) \quad \text{and } X = \ln(a) \tag{8}$$

(Berens, 1984:146)

Linear regression methods are then used to estimate α and β.

Before performing a linear regression on NDI reliability data, the data must be reduced to a set of n pairs, (a_i, p_i), where a_i is the crack length for the ith pair and p_i is the proportion of times the flaw (or flaws) were detected.

Given the n pairs of (a_i, p_i) data points to be fit by the regression analysis, the transformations of equation (8) are performed, resulting in a set of n (X_i, Y_i) pairs.

The variables X and Y are then used in a linear regression analysis resulting in estimates $\hat{\alpha}$ and $\hat{\beta}$ for α and β, respectively. The formulas for $\hat{\alpha}$ and $\hat{\beta}$ are

$$\hat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i - \dfrac{\sum_{i=1}^{n} X_i \cdot \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} X_i^2 - \dfrac{\sum_{i=1}^{n} X_i^2}{n}} \tag{9}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{X} \tag{10}$$

where $\bar{Y}$ and $\bar{X}$ are given by

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}, \quad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{11}$$

(Berens, 1984:147)

The estimated mean $\bar{Y}$ as a function of a is given by

$$\bar{Y}(a) = \hat{\alpha} + \hat{\beta}.\ln(a) \tag{12}$$

(Berens, 1984:147)

The formula for a lower confidence bound on the mean μ_{YL} for a given value is

$$Y_L = \hat{\alpha} + \hat{\beta}.X - t_{(n-2),\gamma}\,(S)\,\sqrt{(1/n) + (X-X)^2/SSX} \tag{13}$$

where

γ is the confidence coefficient

$t(n-2),\gamma$ is the γ th percentile of a t distribution with (n-2) degrees of freedom

$$S = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(Y_i - A - B.X_i)^2} \tag{14}$$

$$SSX = \sum_{i=1}^{n} X_i^2 - \frac{\sum_{i=1}^{n} X_{i\,2}}{n} \tag{15}$$

(Berens, 1984:147)

The inverse Y transformation applied to equation (12) gives the estimate of the POD and, similarly, the inverse Y transform of Y_l gives the confidence bound on the POD function. The equations for the estimate of POD(a) and its lower confidence bound are

$$POD(a) = \frac{\exp(\bar{Y}(a))}{1 + \exp(\bar{Y}(a))} \tag{16}$$

and

$$POD_i(a) = \frac{\exp(\bar{Y}_i(a))}{1 + \exp(\bar{Y}_i(a))} \quad (17)$$

(Berens, 1984:14?)

As discussed in Chapter II, in the event of the observed proportion of detected cracks at a crack length (p_i) being zero or one, then arbitrarily defined values (i.e. 0.01 or 0.999) are assigned, or ,if the sample size is large (i.e. > 30) then the mean estimate is used.

Data Collection

Data sheets for technicians reporting of inspection results were designed for graphic depiction of the flaws. Annex A to the NDI Procedure (Appendix B) shows a blank data form. Data from these individual reports were tallied into the raw data sheet of Appendix C.

Results obtained are then evaluated for statistical significance e.g. goodness of fit, etc, and graphed.

Testing the Assumptions of Regression

On the one hand, regression analysis just fits a line to a set of points. But inferences about the line and the validity of predictions based on the estimated regression require certain assumptions. Thus, it is important in regression analysis to make sure that these assumptions are reasonable and are supported by the data.

Some important assumptions in linear regression can be tested in terms of the residuals. A residual is the

difference between the actual value of y and the value forecast to be the estimated regression line. Thus, if the regression is based on a sample of n data points, or n pairs (x, y), there are n residuals. In linear regression, it is normally assumed that the residuals should be independent and normally distributed with a constant variance (McClave and Benson, 1985:407).

An analysis of the residuals helps decide if the assumptions of linear regression are met. If the residuals are viewed in order, independence suggest that there should not be any systematic patterns. For example, if the residuals tend to be very close to zero for the early data points but much larger for the late data points, the variance may not be constant. If negative residuals cluster together, and the positive residuals do likewise, the independence is unlikely. Some tests are available. The Durbin-Watson test checks for independence, and the chi-square goodness-of-fit test can be used to check for normality (McClave and Benson, 1985:650, 818).

Measures of Model Utility

At the heart of the POD model used for this study is an assumed relationship between the flaw length and its probability of detection (POD), which is represented by transformed linear equation (8). This implied relationship denotes a correlation between the variables.

The Pearson product moment correlation coefficient r,

provides a quantitative measure of the strength of the linear relationship between two variables (McClave and Benson, 1985:418). The closer r is to 1 or -1, the stronger the linear relationship between the variables. However, high correlation does not imply causality.

Another way to measure the contribution of one variable in predicting another is to consider how much the errors of prediction of POD were reduced by using the information of the flaw length. The coefficient of determination, r^2, represents the measure use for this evaluation (McClave and Benson, 1985:422). A measure of r^2 close to 1 indicates a worthwhile relationship.

Chapter Summary

Chapter III specifies the population under evaluation, experimental procedure, and statistical analysis and validation that has to be performed for this reliability study. Based on the guidelines detailed, the following chapter will discuss the findings and analysis of the reliability data collected.

IV. Finding and Analysis

Chapter Overview

This chapter details the field study response, findings and provides analysis results on the data collected from the NDI reliability program.

Field Study Response

Owing to other RAAF work commitments and constraints caused by lack of availability of time and resources (lack of shelf-life magnetic rubber), the Appendix A experimental procedure could not be followed. Out of shelf-life magnetic rubber was used to continue the reliability program. This, in turn, necessitated a change in mix ratios for the magnetic rubber formulation. The NDI Procedure (Appendix B) mixture process reverted to a standard 10cc magnetic rubber base material, with 10 drops of Dibutyl Tin Dilaurate and two drops of Stannous Octoate. This, however, meant that all inspections were performed with reference standards to guarantee the NDI process. Hence, research question 5 was eliminated from this investigation, in favour of the higher precedence research questions' study. Accordingly, a reduced experimental procedure was adopted. This reduced experimental program resulted in a total of 84 inspections being performed on the various test coupons. Tables III and IV show the number of inspections performed on each D6AC test coupon for the flawed and placebo groups, respectively.

TABLE III

Inspections Performed on Flawed Test Coupons

	Coupon Number	Number of Inspections
Flawed Test Coupons	4	4
	53	5
	59	5
	82	4
	145	5
	160	7
	175	7
	272	5
	N	5

TABLE IV

Inspections Performed on Placebo Test Coupons

	Coupon Number	Number of Inspections
Placebo Coupons	43	4
	64	7
	71	5
	91	5
	120	5
	131	7
	136	4

<u>Population Statistics</u>

Although RAAF Base Amberley has 33 staff employed within NDI positions, only 22 are engaged in NDI on a full-time basis. The remainder are supervisors or management staff. Only practicing or trainee NDI technicians are intended to be subject to the field study. The 22 NDI

technicians subjected to this field study had the following composition:

TABLE V

Population Statistics

NDI Classification	
Level 1 Technicians	4
Level 2 Technicians	17
Unspecified	1
Employment History	
Continuous Employment	15
Non-continuous Employment	6
Unspecified	1

Analysis

Table III indicates that the sample size for this research study range from four to seven inspections on each D6AC test coupon, and hence on each flaw size applicable to that test coupon. Given the requirements for sample size from Chapter II, only the test coupons with seven inspections i.e. test coupons serial numbers 160 and 175, were further investigated. This was intended to provide an indication if the POD model would provide useful results.

Accordingly, the information from the raw data sheets for these coupons (Appendix D) was analyzed using the regression functions available with Lotus 123. Appendix D

illustrates the point estimates for the detection probability for the D6AC test coupons. This data was transformed to fit the transformed linear equation (functional form of equation (12)). The equation that resulted was

$$Y = 4.10 + 0.76 \ln(a)$$

The Pearson product moment of correlation, which provides a quantitative measure of the strength of this linear relationship was 0.45. Based on this result, the linear relationship between the flaw size, ln(a), and Y, the transformed POD, can only be classified as moderate. This variation is also indicated by the standard error for the Y estimate being 1.40. Additionally, the value of the coefficient of determination, r_2, was found to be 0.20. This tells us that only 20 percent of the variation among the POD is accounted for by the differences in the flaw size, ln(a). Therefore the model is not particularly useful for making accurate predictions. However, this is more likely to be a result of the effects of the small sample size than from deficiencies with the POD model since the model was tested using the extensive "Have Cracks" data. For example with a sample size of seven, one miss of a flaw can change the detection probability point estimate by 0.13 or 13 percent. Whereas, with a sample size of 30, the point estimate changes only 0.033, or 3.3 percent.

Given the poor predictability of this model, detailed investigation of the influence of some of the human factors

on NDI reliability is not feasible. Additionally, the samples are very small. There were only six NDI technicians without continuous employment in the NDI field. And in both test coupon inspections there was only one NDI technician with non-continuous employment within the NDI field. Thus, research questions 3, and 4 will remain unanswered until data from a larger sample is obtained, and the POD model is better validated.

Nevertheless, further computations were undertaken to glean some useful information, within the constraints of a POD model not being as statistically valid as desired. POD and its associated lower bound confidence level were calculated and the results graphed at Appendixes E and F. The lower bound confidence limit for the POD (functional form of equation (13)) was

$$Y_l = 4.10+0.76.X-1.687.(1.4)\sqrt{(1/40)+(X-\bar{X})^2/50.5}$$

where 1.687 is the 95 percentile of a t distribution with 38 degrees of freedom (i.e. n is 40 obvervations);

1.4 is the standard error of the Y estimate (S); and

50.5 is the value of SSX for the data set used.

From the graph at Appendix F an estimate of the POD at the rated confidence level is 0.17mm (0.007 inch). However, this result has to be considered in light of the POD model utility, i.e. its moderate linear relationship and poor useful for predictions.

The restricted sample size available also prevented

effective analysis of the effect of flaw calls on the overall reliability. While there were flaw calls, see Appendix C, they could not be contrasted to the general population results to establish the changes in the POD function. More data collection is required to allow this analysis to proceed.

Of the 40 flaws on the two D6AC test coupons, there were 26 that were not subject to being missed on any of the seven inspections. However, only four of these flaws were suitable for measurement comparisons. The nature of the flaws on the test pieces meant that the majority were micro cracks, with most linking to their neighbour flaw. As a consequence, most indications did not separately identify each of the flaws, or the NDI technician classified the individual flaws as a composite indication and measured accordingly. Hence, there were only the four candidates for review. No meaningful analysis is possible with such a limited sample. The results of the comparison are provided at Appendix G.

The most limiting factor in analysis of this research is the small sample size as a consequence of the relatively small population. The experimental design should be refined to provide more efficient data collection from a wider population.

Chapter Summary

As a result of a literature review, and experimental

procedures developed from that review, a statistical evaluation of RAAF NDI reliability was undertaken. The results are detailed in this chapter. The analysis of the reliability results using the log logistic model was encouraging, but statistically inconclusive, because of the small sample size. The following chapter will draw conclusions and make recommendations for refining the experimental techniques to allow better validation of the model and more constructive results.

V. Conclusions and Recommendations

Chapter Overview

This chapter reiterates the goals of this program and details the general and specific conclusions that are drawn from the analysis results of the reliability of RAAF NDI. Additionally, recommendations based on both data analysis and observations made in the course of data acquisition are provided.

Program Goals

This program was designed to establish a method of measuring the overall performance of NDI employed within the RAAF. The primary intent was to quantitatively measure the reliability of the Magnetic Rubber method of NDI, as practiced at RAAF Base Amberley, Queensland, to detect cracks in a series of D6AC test coupons. There was also an opportunity to gather and record data on a number of variables associated with the NDI process, human factors, environment, and NDI method.

The basis for evaluation of the RAAF NDI reliability are discussed during the review of available literature and the methodology employed for this study,in Chapters II and III respectively.

Results obtained and the analyses derived from the data are detailed in Chapter IV.

General Conclusions

Because of time and resource constraints, the complete experimental procedure, originally envisaged for this reliability program, was not implemented. However, notwithstanding this limitation, several research objectives were realized. A means of measuring RAAF NDI reliability was established, via the review of available literature. Quantification of flaw detection limits for magnetic rubber inspections at RAAF Base Amberley was then made on the basis of this methodology. However, because of sample size considerations, there was , theoretically at least, increased variability in the POD estimates. Therefore, the results are inconclusive.

Although MIL-A-83444 requires a 90 percent POD with a 95 percent confidence level, there are difficulties with characterizing the NDI reliability on the basis of a single point on the POD function. It does not take into account the variabilty of the overall performance. However, at this point in time, it is still a requirement.

Specific Conclusions

As discussed above, the full experimental procedure was not performed. Hence, many of the specific research questions posed by this study were unanswered, due to lack of sufficient available data. However, some progress was achieved as the primary research question study showed the RAAF Base Amberley NDI technicians were able to detect flaws

of 0.17mm (0.007inch) length, at the specified POD and confidence level. However, this result is derived from a POD function model which had very poor correlation with the flaw length variable. This is possibly due to the statistically small sample size available.

During the course of this study, it was found that the number of D6AC test coupons and hence the number of specific flaw lengths was probably too extensive for the potential statistical accuracy required. The range of flaw lengths above 1.25mm (nearly 50 thousandths of an inch) could be reduced without impinging on experimental integrity.

Cataloging of the flaws was undertaken some time before the actual field study. Further, the flaws were cataloged by an NDI method. No metallurgical evaluation was performed because of the need to re-use the test coupons for other research. Since the flaws were induced into the D6AC test coupons under a corrosively stressed environment, the stress corrosion cracks may still be growing, despite preventative measures being taken. Accordingly, the flaws should be re-cataloged after each base inspection for more accurate base lining of this data.

The scope of data available also limited analysis of the differential between the actual flaw lengths and the reported crack sizes deduced from the magnetic rubber replicas. However, while individual responses may vary, the group findings were not significantly different from the actual flaw lengths.

Lack of data prevented other research questions to be addressed. But, the need for these questions to be analyzed and evaluated is not diminished for this lack.

Recommendations

The logs odds model, validated by Berens, appears to be an acceptable means of specifying the POD function for RAAF NDI reliability evaluations. However, this is based on the limited data collected so far, recognizing that the results obtained are inconclusive. Accordingly, the following recommendations are provided:

1. The RAAF NDI reliability program be continued with data collection from other mainland RAAF bases.
2. The experimental procedure should be amended to reduce the number of test coupons under inspection. Only the D6AC test coupons, serial numbered 53, 82, 175, and 272, should be used in the continuing NDI reliability program.
3. The above nominated test piece be re-cataloged to establish the flaws present and their length (by NDI). Suitable Magnetic rubber replicas of the test coupons be retained as master records.
4. A single person be nominated to oversee the reliability program with the additional task to record results, after each NDI technician trial, and resolve any ambiguity by reference to the master replicas.

5. Selected D6AC test coupons should be the subject of teardown evaluation to accurately determine and fully characterize their flaw content. This would ensure better confidence in the catalog baseline.
6. Based on the data acquired by this further study, more conclusive validation of the POD function should be attempted.
7. Assuming this model validation is successful, the human factors of RAAF NDI should be further investigated.
8. Upon completion of this reliability program phase, other NDI methods, such as eddy current, penetrant, ultrasonics, and radiography, should then be evaluated in a like manner, using appropriate test pieces and procedures.

Chapter Summary

Although some of the thesis research goals were obtained, some were not. Hence, the final chapter of this thesis provides recommendations and conclusions intended to develop a better experimental procedure for the determination of the RAAF NDI reliability.

Appendix A: Experimental Procedure

ROYAL AUSTRALIAN AIR FORCE
NON-DESTRUCTIVE INSPECTION
RELIABILITY PROGRAM
EXPERIMENTAL PROCEDURE

Point of Contact

A single person within the NDISL organization should be appointed co-ordinator for running of this Reliability Program. The NDI Officer, or other OIC NDISL appointee, should be tasked with responsibility for the efficient operation of this experimental procedure.

Reliability Trial

The following steps are to be conducted for each NDI reliability trial undertaken.

1. Select a MRI method. The initial study will investigate MR502 sensitivity and associated NDI Procedure reliability. RAAF Base Amberley staff will also be involved with MR502K sensitivity/reliability studies.

2. Select a Magnetic Field application method. Permanent magnets will be used for the initial study. However, Amberley staff will be used to evaluate DC magnetic field application effects.

3. Assemble the 12 D6AC (with defects) test coupons. Ensure that the test pieces are clean and protectively coated with a corrosion preventative grease.

4. Assemble another 12 similar D6AC coupons (without defects). These test coupons are the control sample. Ensure that the test pieces are clean and protectively coated with a corrosion preventative grease.

5. Randomly mix both groups of D6AC coupons together. Do not mishandle or damage the test pieces during this process.

6. Place the thoroughly mixed test coupons in a numbered sequence.

7. Using the attached Random Number Table haphazardly select a number in the table. Proceeding from this number across the row or down the column (either will do), remove and record four numbers

from the table. Use only the necessary number of digits (one) in each random number to identify the removal interval for each of the 24 D6AC test coupons in their numbered sequence.

8. Beginning at the first D6AC test coupon, step over the number of test coupons indicated by the first random number drawn from the Table and remove the next D6AC test coupon in sequence from the set. Continue this process using the remaining three random number until four test coupons have been selected. These four D6AC test coupons will be given to one NDI technician for the reliability program investigation.

9. Repeat steps 6 to 8 until six randomly selected groups of four D6AC test coupons are selected. Set aside these six sets of test coupons.

10. Randomly select a NDI technician (using a similar method as used for selecting the D6AC test coupon sets) and assign one of the group of four test coupons to him. Repeat until all six sets of test coupons are assigned.

11. Instruct the randomly selected NDI technicians about the purpose of this Reliability program. Stress that the program is designed to establish overall (group) NDI technician reliability. As such, individual results will not be separately identified, and confidentiality of results is guaranteed. The NDI technician identification that is requested, by the NDI Procedure Data Return Sheet, is only to allow contact with the NDI technician should the need arise. For example if there seems to be an anomaly with the test results i.e. test results incorrectly annotated on the data return sheet because of wrong orientation of test pieces, etc. Further, the NDI technician should be advised to perform the task with the same diligence as would 'normally' be used. He should not try to over excel because he is under test.

12. Provide the NDI technician with a copy of the Reliability Program NDI Procedure, as well as a statement of what MRI (i.e. MR502/MR502K) and magnetic field (permanent magnet/electro magnet) are to be used.

13. If the Pre-Inspection Requirements for sensitivity verification of the MRI is successful, only one inspection of each D6AC test coupon for each reliability trial is to be made (assuming the

replicas are not destroyed on removal, or some other catastrophic failure prevents viewing of the potential indications).

14. Each NDI technician is to perform an inspection, in accordance with the NDI Procedure, independent of any outside assistance. Results are to be recorded on the appropriate data sheets, with the consolidated results returned to NDISL.

15. Return all documentation and D6AC test pieces to NDISL.

16. After cleaning the test coupons (if necessary) repeat the above process from step 6 until all NDI technicians have undertaken the program.

Reliability Program Sequence

The RAAF Base Amberley NDI technicians are required to perform the reliability program using:

a. One inspection with MR502 and permanent magnet for field application, then followed by a redundant, yet independent, inspection on the same D6AC test coupons using the same inspection procedure;

b. One reliability trial with MR502 and DC magnetic field application on new D6AC test coupons,

c. One reliability trial with MR502K using permanent magnets and new D6AC test coupons,

d. One reliability trial with MR502 and permanent magnet field application; but performed without doing the Pre-Inspection MRI sensitivity verification called out by the NDI Procedure.

After the Amberley based staff have completed these trials the remaining RAAF NDI technicians can undertake the reliability program, using only MR502 and permanent magnets. RAAF Base Richmond, followed by Williamtown should be the next bases to be tested.

Annex:

A. Random Number Table

RANDOM NUMBER TABLE

Table I Random Numbers

ROW \ COLUMN	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291	90700
2	22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615	99505
3	24130	48360	22527	97265	76393	64809	15179	24830	49340	32081	30680	19655	63348	58629
4	42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758	16379
5	37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263	54613
6	77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394	42880
7	99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
8	96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
9	89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678	44947	05585	56941
10	85475	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593	64952
11	28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610	78188
12	63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703	90322
13	09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613	74952
14	10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104	20285	29975	89868
15	07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372	74461	28551	90707
16	51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
17	02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835	44919	05944	55157
18	01011	54092	33362	94904	31273	04146	18594	29852	71585	85030	51132	01915	92747	64951
19	52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738	17752	35156	35749
20	07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916	19509	25625	58104
21	48663	91245	85828	14346	09172	30168	90229	04734	59193	22178	30421	61666	99904	32812
22	54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524	15227	96909	44592
23	32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012	64161	18296	22851
24	29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10367	07684	36188	18510
25	02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953

Appendix B: NDI Reliability Procedure

RAAF NDI RELIABILITY PROGRAM

MAGNETIC RUBBER INSPECTION (MRI) OF CYCLIC INDUCED CRACKING IN D6AC STEEL TEST COUPONS

Operator Level: All RAAF qualified NDI technicians

Introduction

1. The 'USAF Military Specification on Airplane Damage Tolerance Requirements', MIL-A-83444 assumes and accepts that flaws are inherent in all material. It also asserts that when such materials are used for aircraft production the manufacturer/structures engineer is to ensure that currently available NDI methods are capable of detecting such defects while they are still below the critical size threshold.

2. To enable any manufacturer or engineer to confidently provide critical defect size data they must be privy to the sensitivity of the NDI method and reliability of the NDI technician as individual components and as integrated components of the inspection process. Therefore, this program was introduced to acquire and analysis such as aspects within the RAAF's NDI field.

Purpose

3. This inspection program is designed to provide data:

a. to quantify the minimum, consistently detectable defect size in D6AC steel using MRI methods,

b. to verify the sensitivity of MRI materials, and

c. to establish flaw detection probabilities under a number of operating conditions.

Specific Supervision Requirements

4. The supervisory requirements of DI(AF) TECH 25-11, paragraph 16 will prevail throughout this program.

Applicability

5. The inspection(s) will be undertaken on coupons provided ny NDISL. The NDI technicians required to perform the inspection(s) will be randomly selected by NDISL and will carry out the inspection(s) using a nominated magnetic rubber type, mix ratio, and magnetic field application.

Equipment Requirements

6. The following equipment is required to perform the inspection(s):

a. Magnetic Rubber Inspection Kits:

(1) MR502,

(2) MR502K,

(3) MR502Y,

(4) MR502P.

b. Bell 610 Gaussmeter, complete with transverse probe PN STG1-0404;

c. Electro Magnet PN MRIK49;

d. Electro Magnet PN DA200;

e. Permanent eclipse Magnet PN 813;

f. Permanent eclipse Magnet PN 814;

g. Permanent eclipse Magnet PN 815;

h. Permanent Magnet PN NDISL/MRI/10;

i. 25ml Glass Beakers, Graduated;

j. 50ml Glass Beakers, Graduated;

k. 200ml Glass Beakers, Graduated;

l. 50mm Diameter Petri Dish or similar:

m. ?50mm long 7mm diameter Glass Rod;

n. Lead Tape;

o. 25mm masking tape;

p. Cotton buds;

q. Lint free cloth or substitute;

r. Stereo Microscope WILD M7A (provided by NDISL);

s. Magnifier PN FAA-173;

t. Reference Standard PN NDISL/MRI/11;

u. Solvent General Cleaning MIL-C-38736;

v. "PRESERVAC" or similar protective coating

w. Stopwatch;

x. NDISL D6AC Test Coupons.

Safety Procedures

7. All personnel involved in this inspection program must comply with the safety precautions in DI(AF) AAP 7002.008-1 Section 1 and any other relevant publication.

WARNING

Solvent MIL-C-38736 is a toxic chemical and can cause permanent disability if used without respiratory protection. Further, it must be specifically formulated for use on D6AC steels and must not be used as a general purpose cleaning agent.

8. Caution should be exercised when handling Stannous Octoate. The fluid is capable of serious injury if contact with the eye(s) occurs. Normal eye protection is essential.

Pre-Inspection Requirements

9. Calibrate and balance the gaussmeter and transverse probe combination in accordance with the manufacturer's handbook. Survey all reference standards and test coupons for residual magnetism. Where necessary, demagnetize to the lowest possible level using the MRIK49 electro magnet and suitable pole pieces. The maximum allowable residual magnetism in any item is 4 gauss.

10. Ensure that the MRI base material is not magnetized. This condition, which is a cause of reductions in sensitivity, is evident by the 'clumping' of the ferro magnetic particles and can be eliminated from the material before mixing.

11. Allow the MRI materials to reach ambient temperature before use. A minimum of four hours stabilization is required for 2lb cans of base material and eight hours is required for 7lb cans.

12. An 'on site' test of the magnetic rubber is to be performed on Reference Standard PN NDISL/MRI/11 to certify the sensitivity of the MRI being used. The test must establish which mix ratio gives a clear and accurate indication of the nominated defect in the Reference Standard. This test is to be performed in the following circumstance:

a. at the start of each inspection day,

b. at three hourly intervals until the inspection is completed, and

c. whenever a replacement consumable is introduced into the inspection.

13. The sensitivity verification of the MRI is achieved in the following manner:

a. Prepare a cast identification token which nominates batch number, date and time information. A convenient and practical identification method is by 'mirror writing' on lead tape and affixing this to the Reference Standard outside the Area of Interest (AOI).

b. Form a Dam around the Reference Standard. Tape applied around the Reference Standard edges is adequate for this purpose.

c. Apply a magnetic field of between 25 and 30 gauss to the Reference Standard. This field is measured in air immediately above the AOI and lateral to the Reference Standard. Refer to Figure 1.

d. Prepare the MRI casting material in the ratios prescribed in Table 1. Mixing and pouring is to accomplished as swiftly as possible, commensurate with adequate mixing and minimum aeration.

WARNING

In instance involving the use of MR502K, MR502Y or MR502P, a stopwatch must be used to monitor the elapsed time between addition of the catalyst and filling of the dam.

This time must not exceed 1.5 minutes.
WARNING (Contd)

In instances using MR502, the inspection is to be considered invalid if the time taken between catalyzation and magnetic field application exceeds one half of the estimated pot life.

TABLE I

Magnetic Rubber Mixture Formulae

PN	cc Base	Drops DTD	Drops SO	Drops CS	Pot Life
MR502	20	2	Nil	2	3 mins
MR502K	30	2	2	Nil	2 mins
MR502Y	30	2	2	Nil	2 mins
MR502P	10	3	3	Nil	5 mins

Note: 1. Abbreviations- DTD Dibutyl Tin Dilaurate
SO Stannous Octoate
CS Cure Stabilizer

WARNING

Stannous Octoate catalyst is extremely light sensitive. Ultra Violet (UV) light impinging upon this catalyst will cause it to deteriorate with consequent mix sensitivity loss. This deterioration may still occur with Stannous Octoate stored in light coloured or semi-opaque containers. To minimize UV deterioration, Stannous Octoate should be stored in a light proof container and exposed to ultra violet light sources for a minimum amount of time. Stannous Octoate containers may be rendered light proof by wrapping with light proof tape.

e. Pour the magnetic rubber mix into the prepared dam on the magnetized Reference Standard. Allow this to cure with minimum disturbance.

f. Remove and inspect the cured replica cast under a microscope using a maximum of X15 magnification. If

no crack indications are apparent, carry out another sensitivity verification with a modified mix ratio which increases either pot life and/or particle migration time.

g. Measure the crack indication(s). Documentation accompanying this Reference Standard shows several flaw indications that are numbered and measured. If, as a result of this Pre-Inspection Procedure, a measurement of 0.005 inches cannot be obtained from Indication # 1, additional sensitivity verification inspection(s), with modified mix ratio(s), must be carried out. These inspections are to be carried out with varying pot lives, migration times, cure times and magnetic field strengths until the indication from crack #1 appears as 0.005 inches in length. All sensitivity check details are to be annotated on the proforma included at Annex A.

Inspection Procedure

14. The hole bore and surrounding areas of the D6AC test coupons are to be cleaned and made free of grease, paint/plating flakes and other contaminants. An approved general cleaning solvent such as MIL-C-38736 is to be used. If necessary, paint and plating flakes may be removed by light rubbing with fine abrasive paper or an approved abrasive pad.

15. Using the balanced gaussmeter and probe combination, inspect the test coupon hole bores and their surrounding area for residual magnetic fields. Where necessary demagnetize the coupon to the lowest practical level using the MRIK49 electro magnet or similar. Residual magnetism must not exceed 4 gauss.

16. Mark cast orientation and information on the D6AC test coupon. Orientation is as follows:

a. The top of the D6AC test coupon is in relation to the serial number stamped on the end of the test piece. In any case where orientation is ambiguous, ie test coupons numbered 88, 96, 09 etc, the top surface is assumed to be the surface to which the grip pad is or was fixed.

b. Holes are numbered conventionally, assumed the serial number edge is towards the viewer.

c. Hole reference point of 0/360 degree is the centre of the outboard edge diametrically opposite the numbered edge. Figure 1 refers.

d. "Lateral" and "longitudinal" refers to coupon dimensions, not distances between the holes.

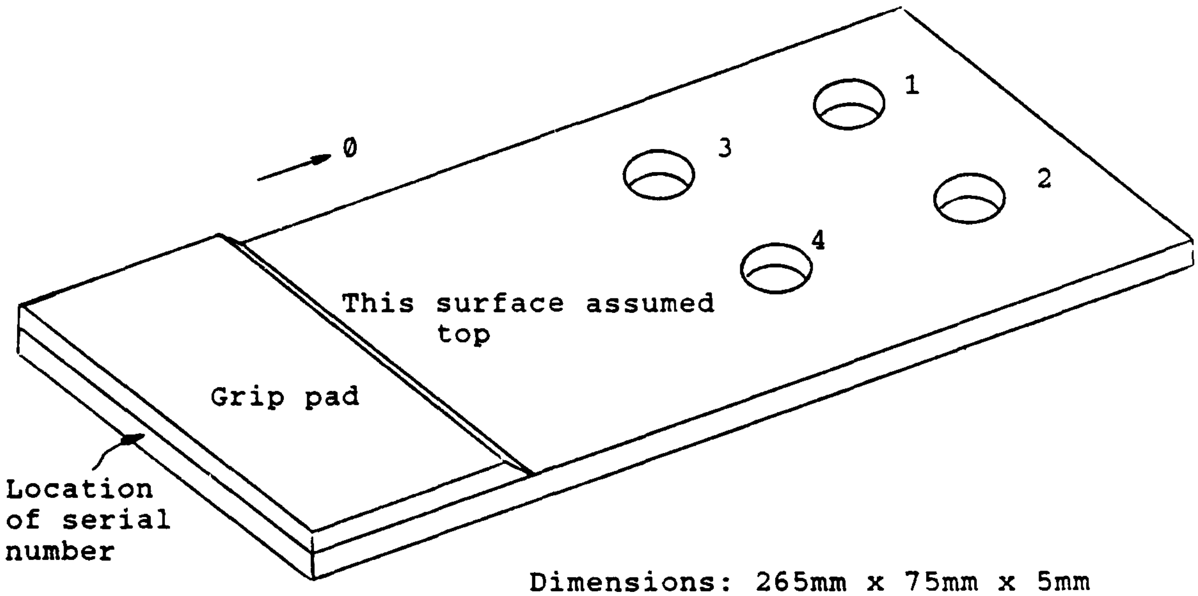


Figure 1. Typical D6AC Test Coupon Showing the Hole Numbering Method and MRI Cast Orientation

17. Decant the required amount of base material. If considered necessary, deaeration is to be carried out before addition of other components. If deaeration is performed, the deaeration method is to be annotated on the Procedure Report (Annex A).

18. Form dams around the holes to be inspected. All holes and surrounding areas may be inspected with one cast. Indication verification or quadrant defect mapping may be carried out on individually dammed holes.

19. Apply a magnetic field to the D6AC coupon under test. A magnetic field strength of 25 to 30 gauss, measured in air in the hole bores, has been found adequate for the indication of defects. The field strength is to be determined and noted before each inspection. Magnetic fields are to be applied by the specified method, ie either permanent or electro magnets (AC and DC).

20. Thoroughly mix the magnetic rubber base material until it contains no streaks or colour variations. This mixing

must be repeated prior to decanting each amount of the base material. This is particularly important when using MR502K, as its lower viscosity results in a faster ferro magnetic particle settling rate.

21. The mix ratio(s) to be employed in producing the magnetic rubber material is that found successful in determining the Pre-Inspection Reference Standard nominated indication. Add the pre-determined quantity of Dibutyl Tin Dilaurate, Stannous Octoate and/or Cure Stabilizer to the base material and mix thoroughly. Ensure thorough mixing of these magnetic rubber components does not infuse the prepared mix with excessive amounts of entrapped air.

22. Fill the prepared dam(s) with the mixed magnetic rubber material. Any excess rubber is to be poured into a petri dish or similar container to assist in monitoring cure times.

23. If necessary, apply further magnetic fields. Duration and orientation of the magnetic fields is to be noted.

24. Cured magnetic rubber replicas are to be removed and inspected for defect indications under magnification not exceeding 15X.

<u>WARNING</u>

A MRI replica is considered to be cured when a 7mm diameter glass rod of 250mm length and 25 grams weight (approx) does not leave an impression in the surface of the cast (or an identically mixed and poured material sample) when dropped from a height of 100mm.

25. All defect indications are to be mapped, measured and catalogued on the inspection data proforma included at Annex A. Surface length measurement of the indications can be either by comparative means (using Reference Standard PN NDISL/MRI/11 as a bench mark) or by direct measurement. The means of measurement has to be annotated at Annex A. Quadrant cracks (extending from the bore hole surface to either the top or bottom flat surface) are to be identified as two cracks (one on each surface) for the purposes of mapping, measurement and cataloging.

<u>Acceptance/Rejection Criteria</u>

26. The absence of any crack indication is not to be

considered a negative result. Reports are to be furnished for every inspection attempted.

Post Inspection Requirements

27. The D6AC test coupons are to demagnetized and cleaned with an approved cleaning solvent. Hole bores and other unprotected areas are to be smeared with a corrosion preventative grease.

Back-Up Inspection Procedure

28. None required.

Reporting Procedure

29. The inspection proforma (Annex A) is to be completed during inspection procedure. The completed inspection data proformae are to be forwarded by Service mail to:

Attention: NDI Officer
No 3 Aircraft Depot (NDISL)
RAAF Base
AMBERLEY QLD 4305

30. Duplicate copies must be held by the inspecting NDI technician's Section until advised by NDISL staff.

Annexes:

A. Data Return Sheet

DATA RETURN SHEET

Inspector Details

NDI Technician Identification
Number (use last four digits of Service Number):

NDI Technician Level Classification (1, 1S, 2, or 3):

Total Number of Years NDI Experience (eg 3.5 years):

Has this NDI Employment been Continuous? (Y or N):
(Have you ever returned fulltime to your
parent mustering duties?)

Present Unit:

Date:

Reliability Test Details

Test Coupon Number: Test Number (1 or 2):

Inspection Performed in Controlled (ie Air
Conditioned) Environment? (Y or N):

Ambient Air Temperature:

MRI Material Type: MR502
MR502K
MR502Y
MR502P

Batch Number:

Magnetic Field Application: Permanent Magnet PN 813
PN 814
PN 815

Electro Magnet PN MRIK49
PN DA200

Reference Standard Results

How many attempts taken to detected the 0.005 inch flaw in the Reference Standard?: _____

Remarks:

Reliability Test Results

	Field Strength/Duration			Flaw Length	
	Lateral	Longitudinal	No. of Cracks	Max	Min
Hole 1					
Hole 2					
Hole 3					
Hole 4					

Measurement Method:

Top Surface Cracks

Hole 1 Hole 2 Hole 3 Hole 4

Crack No.	Length	Crack No.	Length	Crack No.	Length	Crack No.	Length

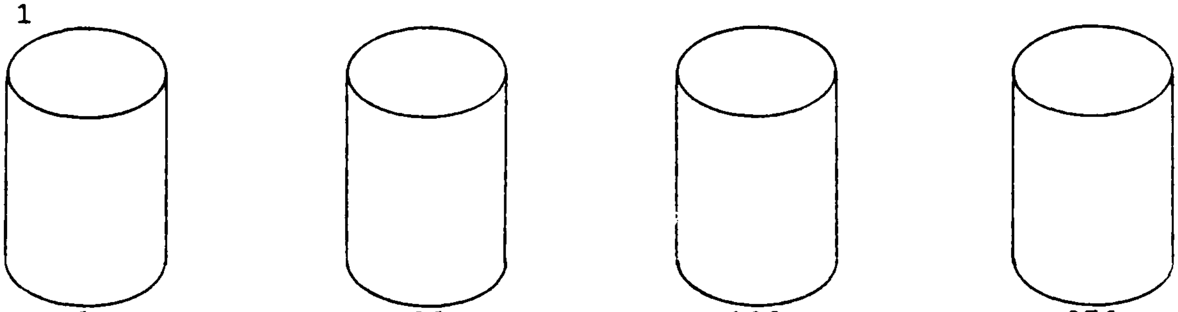
Bottom Surface Cracks

Hole 1 Hole 2 Hole 3 Hole 4

Crack No.	Length	Crack No.	Length	Crack No.	Length	Crack No.	Length

Bore Surface Cracks

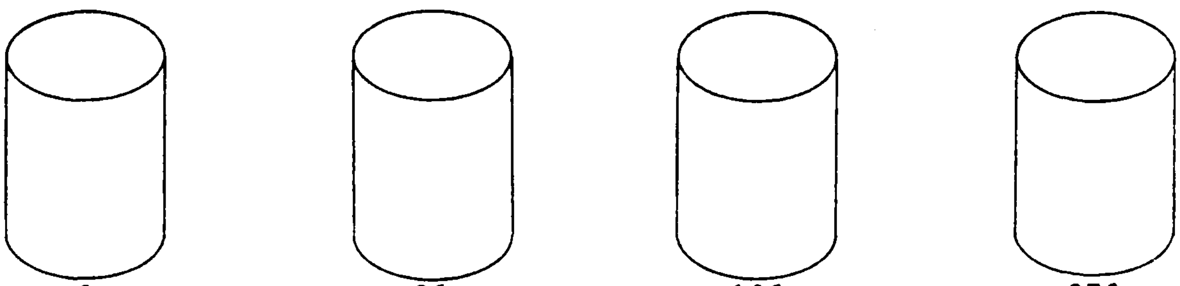
Hole 1



0 90 180 270

Crack No. Length	Crack No. Length	Crack No. Length	Crack No. Length

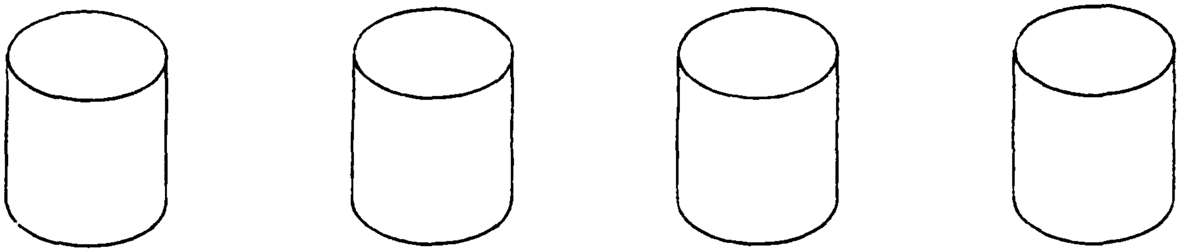
Hole 2



0 90 180 270

Crack No. Length	Crack No. Length	Crack No. Length	Crack No. Length

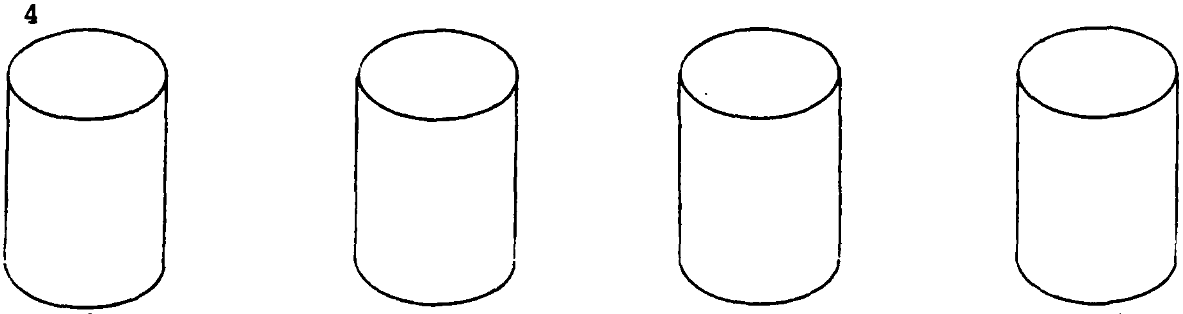
Hole 3



0 90 180 270

Crack No.	Length	Crack No.	Length	Crack No.	Length	Crack No.	Length

Hole 4



0 90 180 270

Crack No.	Length	Crack No.	Length	Crack No.	Length	Crack No.	Length

Additional Comments (if required):

Appendix C: Sample of Raw Data Sheets

RAW DATA SHEET

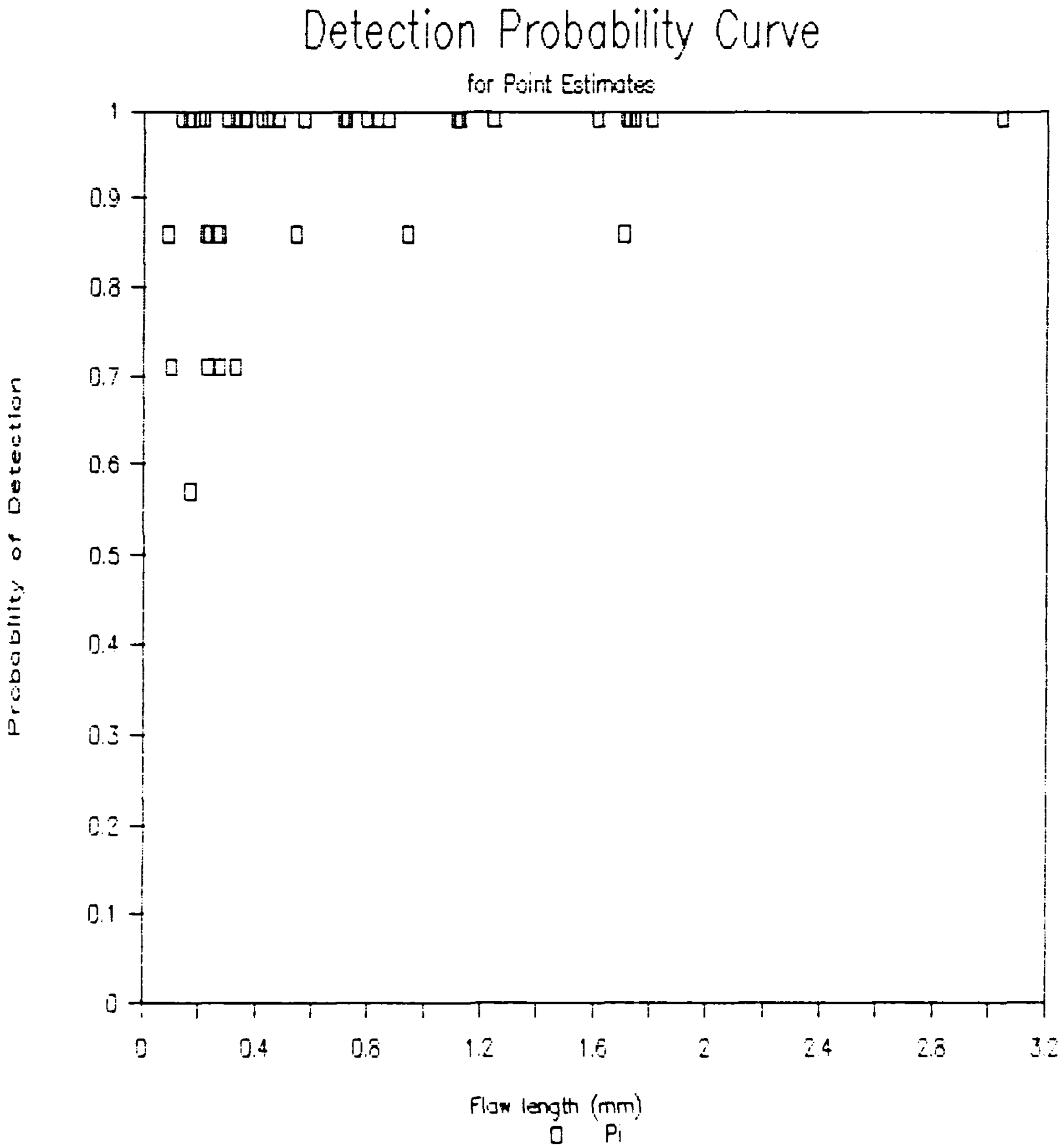
Flaw ID	B	#	Flaw Size	Technician ID 0634	2010	4286	5562	5564	7466	8023	Σ/n	%
160	3	1	0.420	/	/	/	/	/	/	/	7/7	100
160	3	2	0.570	/	/	/	/	/	/	/	7/7	100
160	3	3	0.330	/	/	/	/	/	/	/	7/7	100
160	3	4	1.120	/	/	/	/	/	/	/	7/7	100
160	3	5	0.100	/	0	/	/	0	/	/	5/7	71
160	3	6	0.330	/	/	/	/	/	/	/	7/7	100
160	3	7	1.240	/	/	/	/	/	/	/	7/7	100
160	3	8	1.110	/	/	/	/	/	/	/	7/7	100
160	3	9	0.300	/	/	/	/	/	/	/	7/7	100
160	3	10	0.200	/	/	/	/	/	/	/	7/7	100
160	3	11	1.080	/	/	/	/	/	/	/	7/7	100
160	3	12	0.160	/	/	/	/	/	/	/	7/7	100
160	4	1	0.230	0	/	/	/	/	/	/	6/7	86
160	4	2	0.790	/	/	/	/	/	/	/	7/7	100
160	4	3	0.440	/	/	/	/	/	/	/	7/7	100
160	4	4	0.330	/	0	/	/	/	/	0	5/7	71
160	4	5	1.740	/	/	/	/	/	/	/	7/7	100
160	4	6	0.720	/	/	/	/	/	/	/	7/7	100
160	4	7	0.870	/	/	/	/	/	/	/	7/7	100
160	4	8	0.360	/	/	/	/	/	/	/	7/7	100
160	4	9	0.480	/	/	/	/	/	/	/	7/7	100

160	4	10	1.610	/	/	/	/	/	/	/	7/7	100
160	4	11	1.710	/	/	/	/	/	/	/	7/7	100
F				5	4	6	4	–	4	–		
Σ/n				5/23	4/23	6/23	4/23	0/23	4/23	0/23		
%				21.7	17.4	26.1	17.4	0	17.4	0		

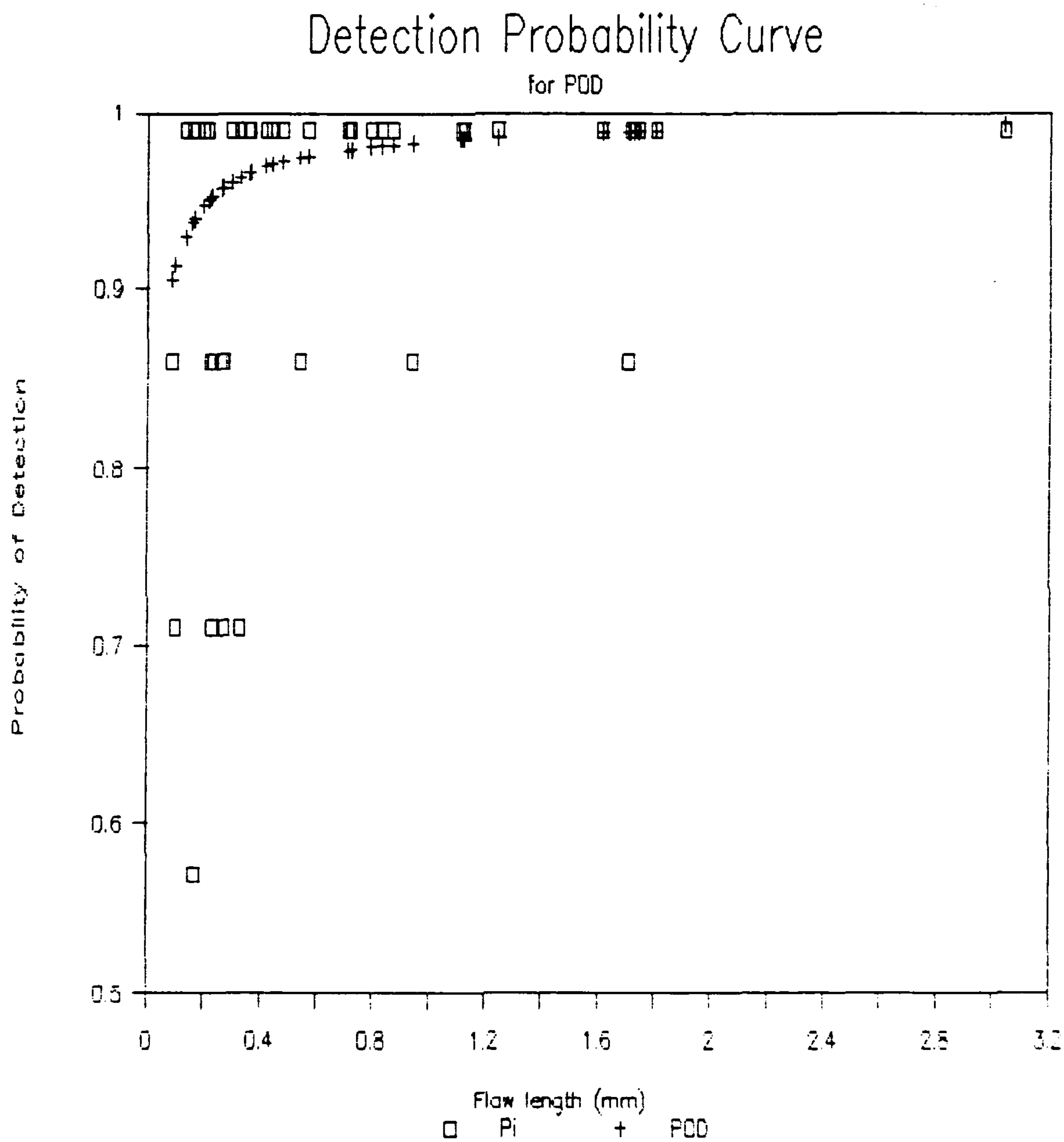
RAW DATA SHEET

Flaw ID ID	B	#	Flaw Size	Technician ID 0634	2010	2286	5562	5562	7266	5023	Σ/n	%
175	1	1	0.168	1	0	0	1	0	1	1	4/7	57
175	1	2	0.216	1	1	1	1	1	1	1	7/7	100
175	1	3	0.540	1	1	1	0	1	1	1	6/7	86
175	1	4	0.087	0	1	1	1	1	1	1	6/7	86
175	1	5	0.364	1	1	1	1	1	1	1	7/7	100
175	3	1	0.271	1	0	1	1	0	1	1	5/7	71
175	3	2	1.723	1	1	1	1	1	1	1	7/7	100
175	3	3	0.709	1	1	1	1	1	1	1	7/7	100
175	3	4	0.270	1	0	1	1	1	1	1	6/7	100
175	4	1	0.229	1	0	1	0	1	1	1	5/7	71
175	4	2	0.223	1	0	1	1	1	1	1	6/7	86
175	4	3	0.263	1	0	1	1	1	1	1	6/7	86
175	4	4	1.703	1	0	1	1	1	1	1	6/7	86
175	4	5	0.939	1	0	1	1	1	1	1	6/7	86
175	4	6	3.041	1	1	1	1	1	1	1	7/7	100
175	4	7	0.831	1	1	1	1	1	1	1	7/7	100
175	4	8	0.135	1	1	1	1	1	1	1	7/7	100
F				–	–	4	–	5	–	–		
Σ/n				0/17	0/17	4/17	0/17	5/17	0/17	0/17		
%				0	0	235	0	294	0	0		

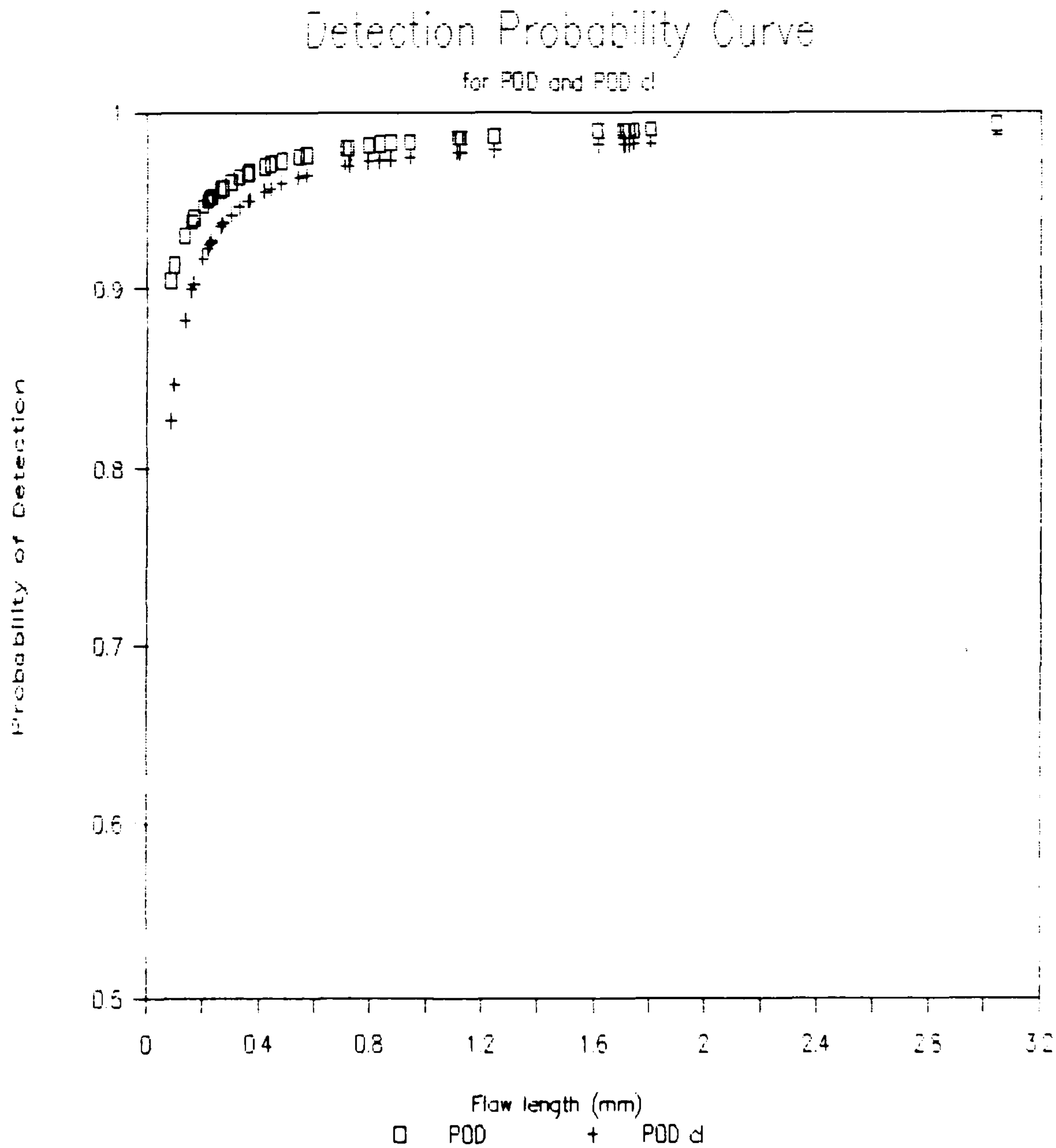
Appendix D: <u>Point Estimate Detection Probability Curve</u>



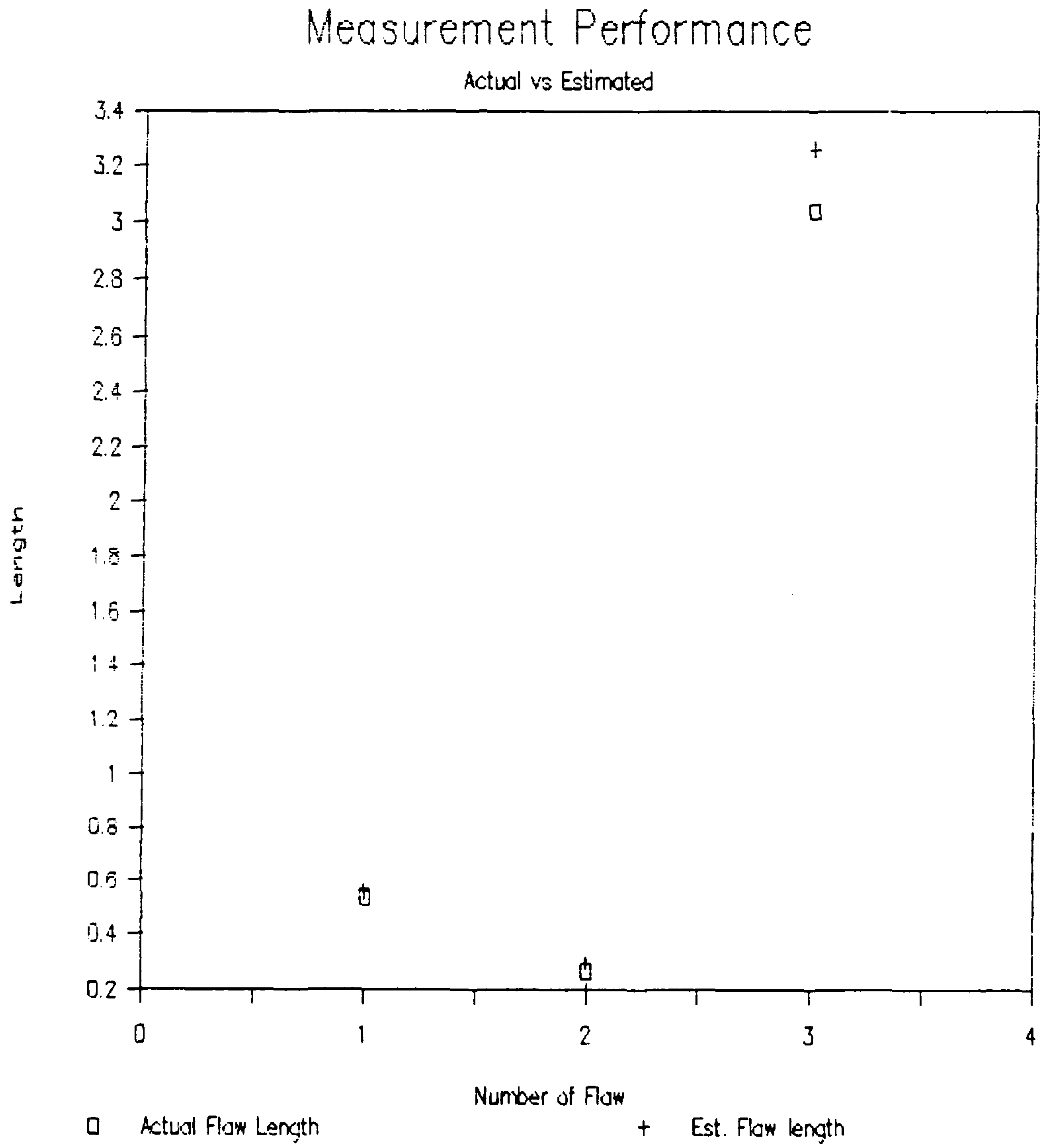
Appendix E: Detection Probability Curve



Appendix F: Detection Probability Curve
POD and Confidence Limit



Appendix G: Measured vs Actual Flaw Length Graph



Bibliography

Berens, A. P. and P. W. Hovey. Flaw Detection Reliability Criteria, Volume 1 - Methods and Results. Technical Report AFWAL-TR-84-4022. Materials Laboratory, Air Force Wright Aeronautical Laboratories, Air Force System Command, Wright-Patterson Air Force Base, Ohio, April 1984 (AD A142 001).

_____. Evaluation of NDE Reliability Characterization. Technical Report AFWAL-TR-81-4160. Materials Laboratory, Air Force Wright Aeronautical Laboratories, Air Force System Command, Wright-Patterson Air Force Base, Ohio, December 1981 (AD-A114-467).

Chin Quan, H. R. and I. G. Scott. "Operator Performance and Reliability in NDI", Research Techniques in Non-Destructive Testing, Sharpe, R. S. (Ed) Volume 3:323-354, Academic Press, London, 1977.

Headquarters, US Army Material Command. Engineering Design Handbook, Development Guide for Reliability, Part IV Reliability Measurement. Technical Report AMCP-707-198. Letterkenny Army Depot, Chambersburg, Pennsyvania, January 1976 (AD-A027 371).

Lewis, W. H., B. D. Dodd, W. H. Sproat, and J. M. Hamiliton. Reliability of Nondestructive Inspection. Technical Report SA-ALC/MME-76-6-38-1. San Antonio Air Logistics Center, Kelly Air Force Base, Texas, December 1978 (AD-A072 097).

Lund, Robert T., Floyd R. Tuler, and John R. Elliott. Life Forecasting as a Logistics Technique. Technical Report AMMRC TR 82-2. Army Materials and Mechanics Research Center, Watertown, Massachusetts, January 1982 (AD-A114 630).

McClave, James T. and P. G. Benson. Statistics for Business and Economics. Dellen Publishing Company, San Francisco, 1985.

Military Standard Non Destructive Inspection, Qualification and Certification of Personnel MIL-STD-410D, USAF, July 1979.

Military Standard Aircraft Structural Integrity Programme, Airplane Requirements MIL-STD-1530A(11), USAF, December 1975.

Packman, P. F. "Reliability of Flaw Detection by Nondestructive Inspection", Nondestructive Inspection and Quality Control, Metals Handbook, Volume 11:414-424, 1976.

Recommended Practice for Demonstration of Non-Destructive Evaluation (NDE) Reliability on Aircraft Production Parts, ASNT, Columbus, Ohio, 1976.

Rummel, Ward D. The Detection of Fatigue Cracks by Non Destructive Testing Methods. Technical Report NASA CR-2369, 1974.

Rummel, Ward D., Steven J. Mullen, Brent K. Christner, Frank B. Ross, and Robert E. Muthart. Reliability of Nondestructive Inspection (NDI) of Aircraft Engine Components. Technical Report SA-ALC/MM-8151. San Antonio Air Logistics Command, Kelly Air Force Base, Texas, January 1984 (AD-A155 320).

Stone D. E. W. Non-destructive Inspection and the Implementation of a Damage Tolerant Design Philosophy. Technical Memorandum Structures 982. Royal Aircraft Establishment, Farnborough, Hants, United Kingdom, February 1981 (AD-A102 867).

VITA

Squadron Leader Mark Cassidy was born on 15 July 1954 in Miles, Queensland. He completed his secondary education at St Edmund's College in Ipswich, Queensland, and enlisted in the Royal Australian Air Force (RAAF) on 22 January 1973 as an Officer Cadet with the Diploma Cadet Squadron (DCS), RAAF Base Frognall, Melbourne. He attended the Royal Melbourne Institute of Technology (RMIT), from which he received a Diploma of Aeronautical Engineering. Upon graduation in November 1976, he was appointed to the Engineering Branch of the RAAF in the Aeronautical Category. Since then he has held a number of staff, unit, and depot positions. Before entering the School of Systems and Logistics, Air Force Institute of Techology (AFIT), in June 1986, he was the Officer-in-Charge of the Non-Destructive Inspection Standards Laboratory, RAAF Base Amberley, Queensland. After graduation from AFIT, Squadron Leader Cassidy will be employed within Capital Projects Division of the Logistics Branch, Headquarters Support Command, Melbourne. He is married to Susan Patricia Bollard, a high school teacher.

Permanent address: 4 Albany Place
Mount Martha
Victoria. 3934.
Australia

4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFIT/GLM /LSMA/87S-11	5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION School of Systems and Logistics	6b. OFFICE SYMBOL *(If applicable)* AFIT/LSM	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS *(City, State, and ZIP Code)* Air Force Institute of Technology Wright-Patterson AFB OH 45433-6583		7b. ADDRESS *(City, State, and ZIP Code)*
8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL *(If applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS *(City, State, and ZIP Code)*	10. SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*

See Box 19

12. PERSONAL AUTHOR(S)
Mark Cassidy, Dip Eng (Aero), SQNLDR (RAAF)

13a. TYPE OF REPORT MS Thesis	13b. TIME COVERED FROM ______ TO ______	14. DATE OF REPORT *(Year, Month, Day)* 1987 September	15. PAGE COUNT 94

16. SUPPLEMENTARY NOTATION

17. COSATI CODES	18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*

Block 19

Abstract

The purpose of this research was to establish, via examination of the available literature, an appropriate means of quantifying the reliability of Non-Destructive Inspection (NDI) as practised by the Royal Australian Air Force (RAAF) NDI technicians. Further, actual measurement of this NDI reliability was to be attempted and the correlation, if any, between the NDI technician's reported and measured results and the actual flaw lengths was to be established.

Apart from producing crack size detectibility curves several human factors of the NDI process were to be investigated as part of this research. Influences of personnel variables are considered important. This study was designed to evaluate the effects on NDI reliability on whether or not the technicians, employment has been continuous within the NDI trade; if there is any correlation between experience level and the reliability results obtained; and, the effect of false calls. The effectiveness of reference standards called for by the NDI Procedure was also to be the subject of review.

This study was, unfortunately, constrained by time and lack of resources. Hence, to achieve results the experimental design was modified, with a subsequent effect on the data collection and ability to investigate some of the research questions.

This study found that the log logistic model was an acceptable Probability of Detection (POD), based on other recent research efforts. However, analysis of reliability results using this model were encouraging, but statistically inconclusive, because of the small sample size available.

Among the recommendations provided are suggestions to improve the experimental procedure, expand the sample size, and continue reliability data collection and analysis to better validate the POD model and answer the research questions made by this study.